

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 05:53:59 ; Search time 3768 Seconds  
(without alignment)  
10943.977 Million cell updates/sec

Title: US-09-923-236-1

Perfect score: 1008  
Sequence: 1 agacagactaaaagccat.....tagataatcagaccatg 1008

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:  
1: gb Da: \*  
2: gb Hcg: \*  
3: gb In: \*  
4: gb On: \*  
5: gb Ov: \*  
6: gb Pat: \*  
7: gb Ph: \*  
8: gb Pl: \*  
9: gb Pr: \*  
10: gb Ro: \*  
11: gb Sts: \*  
12: gb Sy: \*  
13: gb Un: \*  
14: gb Vi: \*  
15: em Ba: \*  
16: em Fun: \*  
17: em Hum: \*  
18: em In: \*  
19: em Mu: \*  
20: em Om: \*  
21: em Or: \*  
22: em Ov: \*  
23: em Pat: \*  
24: em Ph: \*  
25: em Pl: \*  
26: em Ro: \*  
27: em Sts: \*  
28: em Un: \*  
29: em Vi: \*  
30: em Hcg Hum: \*  
31: em Hcg Inv: \*  
32: em Hcg Other: \*  
33: em Hcg Mus: \*  
34: em Hcg Pln: \*  
35: em Hcg Rod: \*  
36: em Hcg Mam: \*  
37: em Hcg Vrt: \*  
38: em Gy: \*  
39: em Htgo Hum: \*  
40: em Htgo Mus: \*  
41: em Htgo Other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1008	100.0	1008	6 AR263643	AR263643 Sequence
2	997	98.9	998	6 AX061225	AX061225 Sequence
3	981.6	97.4	1325	6 BD131596	BD131596 Secretory
4	578.4	57.4	96481	9 AC104811	AC104811 Homo sapi
5	575.2	57.1	157836	2 AC060228	AC060228 Homo sapi
6	558	55.4	639	6 BD109282	BD109282 EST and e
7	537.6	53.3	185969	2 AC069037	AC069037 Homo sapi
8	470	46.6	544	6 BD107961	BD107961 EST and e
9	460	45.6	657	6 AR263644	AR263644 Sequence
10	449	44.5	470	6 BD071746	BD071746 Secreted
11	421	41.8	450	6 BD071795	BD071795 Secreted
12	419.4	41.6	185969	2 AC069037	AC069037 Homo sapi
13	339.6	33.7	485	6 BD122986	BD122986 EST and e
14	286.4	28.4	296	6 BD071681	BD071681 Secreted
15	122.6	12.2	259038	2 AC099443	AC099443 Rattus no
16	121.2	12.0	3489	6 AR072952	AR072952 Sequence
17	121.2	12.0	3489	6 AR254713	AR254713 Sequence
18	121.2	12.0	3489	6 AR261360	AR261360 Sequence
19	121.2	12.0	3489	6 AX107938	AX107938 Sequence
20	121.2	12.0	3489	14 KSUS2064	US2064 Kaposi's ba
21	121.2	12.0	3207	6 AR065852	AR065852 Sequence
22	121.2	12.0	3207	6 AR127850	AR127850 Sequence
23	121.2	12.0	3207	6 AR194752	AR194752 Sequence
24	121.2	12.0	137508	14 KSUT5698	US5698 Kaposi's ba
25	117.8	11.7	28559	14 AF148805	AF148805 Kaposi's
26	116.6	11.6	1365	10 BC049645	BC049645 Mus muscu
27	115.2	11.4	133661	14 U93872	U93872 Kaposi's ba
28	113.2	11.2	158249	2 AL953878	AL953878 Dantio ter
29	111.4	11.1	293592	2 AC094571	AC094571 Rattus no
30	111	11.0	177100	2 AC112388	AC112388 Rattus no
31	111	11.0	246774	2 AC095078	AC095078 Rattus no
32	110.8	11.0	2931	14 AP192756	AF192756 Kaposi's
33	110.6	11.0	5719	10 MCCT851	AF051726 Mus muscu
34	110.6	11.0	188860	10 AC092203	AC092203 Mus muscu
35	110.4	11.0	2286	6 AX460942	AX460942 Sequence
36	110.4	11.0	168210	2 AC116700	AC116700 Mus muscu
37	109.6	10.9	224198	2 AC106674	AC106674 Rattus no
38	109.6	10.9	249541	2 AC095672	AC095672 Rattus no
39	109.6	10.9	249774	2 AC126196	AC126196 Rattus no
40	109.4	10.9	272545	2 AC090533	AC090533 Mus muscu
41	109	10.8	182534	10 AC074046	AC074046 Mus muscu
42	108	10.7	186559	10 AL606742	AL606742 Mouse DNA
43	108	10.7	180944	2 AC113108	AC113108 Mus muscu
44	107.8	10.7	178597	10 AL807824	AL807824 Mouse DNA
45	107.8	10.7	238518	2 AC132752	AC132752 Rattus no

## ALIGNMENTS

RESULT 1  
AR263643  
LOCUS AR263643 1008 bp DNA  
DEFINITION Sequence 1 from patent US 6331413.1. linear PAT 29-JAN-2003  
ACCESSION AR263643  
VERSION AR263643.1 GI:28075576  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1008)  
AUTHORS Adler,D.A. and Sheppard,P.O.  
TITLE Secreted salivary Zsieg3 Polypeptide  
JOURNAL Patent: US 6331413-A 1 18-DEC-2001;  
FEATURES Location/Qualifiers



Db 541 CAGCCACACCTGTAGAGCTGAGCTGCTGAGAGGAGCCCTGTGTGAGCTGAGCTGCTG 600  
 QY 606 CAGAGCACCCTGTGAGCTGAGCTGCTGAGAGGAGCAGCTGTGTGAGCTGAGCTGCTG 665  
 Db 601 CAGAGCACCCTGTGAGCTGAGCTGCTGAGAGGAGCAGCTGTGTGAGCTGAGCTGCTG 660  
 QY 666 CAGAGCACCCTGTGAGCTGAGCTGCTGAGAGGAGCAGCTGTGTGAGCTGAGCTGCTG 725  
 Db 661 CAGAGCACCCTGTGAGCTGAGCTGCTGAGAGGAGCAGCTGTGTGAGCTGAGCTGCTG 720  
 QY 726 CAGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 785  
 Db 721 CAGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 QY 786 GAAATCTCTAGAGAGTACATGAGTGTATCTATCTATCTATCTATCTATCTATCTATCT 845  
 Db 781 GAAATCTCTAGAGAGTACATGAGTGTATCTATCTATCTATCTATCTATCTATCTATCT 840  
 QY 846 CTACAAAGCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 905  
 Db 841 CTACAAAGCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 900  
 QY 906 TCTCACTACATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 965  
 Db 901 TCTCACTACATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 960  
 QY 966 CATTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1002  
 Db 961 CATTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 997  
  
 RESULT 3  
 BD131596 1325 bp DNA linear PAT 18-SEP-2002  
 LOCUS  
 DEFINITION Secretary proteins and polynucleotides encoding the same.  
 ACCESSION BD131596 GI:23226541  
 VERSION BD131596.1  
 KEYWORDS JP 2002503634-A/5.  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.  
 1 (bases 1 to 1325)  
 Jacoby, K., McCoy, J.M., Lavallie, B.R., Racie, L.A., Treac, M.,  
 Spaulding, V. and Agostino, M.J.  
 Secretory proteins and polynucleotides encoding the same  
 Patent: JP 2002503634-A 5 05-FEB-2002;  
 GENETICS INSTITUTE INC  
 OS Unidentified  
 PN JP 2002503634-A/5  
 PD 05-FEB-2002  
 PE 01-JUL-1998 JP 2000501179  
 PR 02-JUL-1997 US 08/867195 27-OCT-1997 US 08/958304 PI  
 KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
 MAURICE TREACY,  
 PI VIKKI SPAULDING, MICHAEL J AGOSTINO  
 PC C07K14/47, A61K38/00, A61P3/10, A61P7/00, A61P7/02, A61P7/04 PC  
 A61P7/06, A61P11/06,  
 PC A61P17/02, A61P19/02, A61P19/04, A61P19/10, A61P21/00, A61P21/04,  
 PC A61P25/00,  
 PC A61P25/14, A61P25/16, A61P25/28, A61P29/00, A61P31/04, A61P31/10,  
 PC A61P31/12,  
 PC A61P33/00, A61P35/00, A61P35/04, A61P37/04, A61P37/06, A61P37/08,  
 PC A61P43/00,  
 PC C12N5/10, C12N15/09, C12P21/02, A61K39/395, A61K39/395, A61K37/02,  
 PC C12N5/00,  
 PC C12N15/00  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 CC Secretory proteins and polynucleotides encoding the same FH  
 Key location/Qualifiers  
 FT source 1..1325  
 /organism='Unidentified',  
 location/Qualifiers

source 1..1325  
 /organism='unidentified'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:32644'  
 BASE COUNT 407 a 289 c 241 g 388 t  
 ORIGIN  
 Query Match 97.4%; Score 981.6; DB 6; Length 1325;  
 Best Local Similarity 99.5%; Pred. No. 3.6e-189;  
 Matches 995; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 10 AAAAAAGCAGATGATCTTTGTTCTCTTAAAGAAAGAAATATATTTAAAAATAC 69  
 Db 9 AAAAAAGCAGATGATCTTTGTTCTCTTAAAGAAAGAAATATATTTAAAAATAC 68  
 QY 70 ATTGCGATTTTCTTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 129  
 Db 69 ATTGCGATTTTCTTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 128  
 QY 130 GAAGCTTCTCTTTGGGCTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 189  
 Db 129 GAAGCTTCTCTTTGGGCTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 188  
 QY 190 CTTCAATGATGAGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 249  
 Db 189 CTTCAATGATGAGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 248  
 QY 250 TTATGCGATGAGAAATTTACACCTCTCTTTATTAATGCGCAGATGAAATACAGTCCCA 309  
 Db 249 TTATGCGATGAGAAATTTACACCTCTCTTTATTAATGCGCAGATGAAATACAGTCCCA 308  
 QY 310 TTACCTGGAATATCTTACATGACACAGGGTTACCTTGATCCCTGGAATCTTAACTTC 369  
 Db 309 TTACCTGGAATATCTTACATGACACAGGGTTACCTTGATCCCTGGAATCTTAACTTC 368  
 QY 370 TCTGGAATTCCTTAATGATCTATCAATCGTGTGTTTCCCTTAAGTACTGAGTTGAATGT 429  
 Db 369 TCTGGAATTCCTTAATGATCTATCAATCGTGTGTTTCCCTTAAGTACTGAGTTGAATGT 428  
 QY 430 TCTCTCTCTCTCTCTGAGGGGTTCCCGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 489  
 Db 429 TCTCTCTCTCTCTCTGAGGGGTTCCCGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 488  
 QY 490 TGCAGCACCGCTGCGCCCACTTAATGACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 549  
 Db 489 TGCAGCACCGCTGCGCCCACTTAATGACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 548  
 QY 550 CACACCTGTAGCAGCTGAGCTGCTGACAGGGGCTCTGTGACGTGAGCTGCTGACAG 609  
 Db 549 CACACCTGTAGCAGCTGAGCTGCTGACAGGGGCTCTGTGACGTGAGCTGCTGACAG 608  
 QY 610 GGCACCTGTTGAGACATGAGCTGCTGACAGAGGACCTGTTGACGTGAGCTGCTGACAG 669  
 Db 609 GGCACCTGTTGAGACATGAGCTGCTGACAGAGGACCTGTTGACGTGAGCTGCTGACAG 668  
 QY 670 GGCACCTGTTGAGACATGAGCTGCTGACAGAGGACCTGTTGACGTGAGCTGCTGACAG 729  
 Db 669 GGCACCTGTTGAGACATGAGCTGCTGACAGAGGACCTGTTGACGTGAGCTGCTGACAG 728  
 QY 730 CAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789  
 Db 729 CAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 788  
 QY 790 TTCTCTGAAAGATGACATGAGGTTCAATTTCTATCTATGATGAGAAATTAAGAAATCTAC 849  
 Db 789 TTCTCTGAAAGATGACATGAGGTTCAATTTCTATCTATGATGAGAAATTAAGAAATCTAC 848  
 QY 850 AAAAGTTTCTTTCTTTTCAAGATATTTCAATCTGTTGTTGTTGTTGTTGTTGTTGTTG 909  
 Db 849 AAAAGTTTCTTTCTTTTCAAGATATTTCAATCTGTTGTTGTTGTTGTTGTTGTTGTTG 908  
 QY 910 ACTACATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 968



```

repeat_region      /rpc_family="L1"      8841_..9197
repeat_region      /rpc_family="L1"      9198_..9354
repeat_region      /rpc_family="L1"      9347_..9541
repeat_region      /rpc_family="L1"      9562_..9589
repeat_region      /rpc_family="AT_rich"  10635_..10656
repeat_region      /rpc_family="AT_rich"  10875_..10895
repeat_region      /rpc_family="AT_rich"  10916_..11033
repeat_region      /rpc_family="L2"      12056_..12093
repeat_region      /rpc_family="AT_rich"  12401_..12394
repeat_region      /rpc_family="L1"      13404_..13440
repeat_region      /rpc_family="AT_rich"  16713_..16842
repeat_region      /rpc_family="MER1_type" 17320_..18195
repeat_region      /rpc_family="L2"      19746_..19843
repeat_region      /rpc_family="L1"      19879_..20155
repeat_region      /rpc_family="MALR"    20172_..20325
repeat_region      /rpc_family="r_rich"   20244_..21091
repeat_region      /rpc_family="L1"      21093_..21382
repeat_region      /rpc_family="A1u"     21395_..21585
repeat_region      /rpc_family="A1u"     21652_..21749
repeat_region      /rpc_family="A1u"     22498_..23093
repeat_region      /rpc_family="L1"      23070_..24215
repeat_region      /rpc_family="L1"      24365_..26232
repeat_region      /rpc_family="L1"      26227_..27201
repeat_region      /rpc_family="L1"      27202_..27243
repeat_region      /rpc_family="rRNA-Lys-MAG" 27244_..27772
repeat_region      /rpc_family="L1"      27799_..27854
repeat_region      /rpc_family="(TA)n"    27877_..27933
repeat_region      /rpc_family="L1"      28004_..28057
repeat_region      /rpc_family="AT_rich"  28778_..28817
repeat_region      /rpc_family="L2"      29000_..29046
repeat_region      /rpc_family="AT_rich"  29249_..29285
repeat_region      /rpc_family="AT_rich"

```

Query Match	57.4%	Score 578.4;	DB 9;	Length 96481;
Best Local Similarity	89.1%;	Pred. No. 4.8e-107;		
Matches 624;	Conservative	0;	Mismatches 76;	Indels 0;
			Gaps	0;

Qy 199 TGAGGATGACAAATGAGATGGTACCCAGCTTCAATCCATCTGAAATATTCCTAATGGCAT 258  
Db 60879 TTAGAGTGAATACATGATGGTACCCAGCTTCAATCCATCTCTGAATATTCCTTAATGGCAT 60933  
Qy 259 ACGGAATTTAACCACTCTCTTTATTATTATGCGCCAGTGAATATACAGTCCCCAGTTACCTGG 318

Db	60939	ACGGAAATTTTACCACTCTCTCTTTATTATGCCCAGGAAATACAGTCCCAAGTTACCTCTG	60999
OY	319	GAATACTTACACTGACACAGGGTTACCTTGGTATCCCTGGAATTTTAACTTCTCTGGATT	378
Db	60999	GAATACTTACACTGACACAGGGTTACCTTGGTATCCCTGGAATTTTAACTTCTCTGGATT	61058
OY	379	CCCTATGCTATCACATCCGAGGTTTTCCCTTAACTGATCAGTGAATGTTCCCTCTG	438
Db	61059	CCCTATGCTATCACATCCGAGGTTTTCCCTTAACTGATCAGTGAATGTTCTCTCTG	61118
OY	439	CCCTCTAAGGGGTTTTCCCTTGTCCCTCTTCAAGTTTTTTTTTTCAGACGTGCAGACC	498
Db	61119	CCCTCTAAGGGGTTTTCCCTTGTCCCTCTTCAAGTTTTTTTTTTCAGACGTGCAGACC	61178
OY	499	CGCGGCCCCACCTAATTGACGTGAGCTGCTGAGCTGACCTTACAGCCACCTGT	558
Db	61179	CGCTGCCCACTAATTGACGTGAGCTGCTGAGCTGACCTTACAGCCACCTGT	61238
OY	559	AGCAGCTGAGCCTGTGTCAGAGGGGCCCTGTGTGACGTGAGCCTGTGACAGGCACTGT	618
Db	61239	AGCAGCTGAGCCTGTGTCAGAGGGGCCCTGTGTGACGTGAGCCTGTGACAGGCACTGT	61298
OY	619	TGAGAGTGAGCCTGCTGCAAGGCACTGTGTGACGTGAGCCGTGCGCAGAGGCACTGT	678
Db	61299	TGAGAGTGAGCCTGCTGCAAGGCACTGTGTGACGTGAGCCGTGCGCAGAGGCACTGT	61358
OY	679	TGAGTGAGAGCCAGCTGCAGAGGAACCTTACCAAGCTGAGCCTGTCTACAGCCAGCCTGC	738
Db	61359	TGAGTGAGAGCCAGCTGCAGAGGAACCTTACCAAGCTGAGCCTGTCTACAGCCAGCCTGC	61418
OY	739	TGCCCCAGAACCTCACACCTTCTCCCTCTTGAACAGGATGTTTATATCTTACAC	798
Db	61419	TGCCCCAGAACCTCACACCTTCTCCCTCTTGAACAGGATGTTTATATCTTACAC	61478
OY	799	AGAGTACCATGGGTTCAATTTCTATACGTATGCAGAATATAGGAATCTTACAAAAGTTTT	858
Db	61479	TATATATGATGAGAAATGAGATTGTGAGAAAGGGGAAAAAAAACACCTCCCAAGCAA	61538
OY	859	CTTCTCTTTCCAAAGACTATTTTCATTCCTCTGATTTTACAA	898
Db	61539	ATCTATATCACTTAAATATTTTGAACAGCTTATTTACAA	61578

RESULT	5
LOCUS	AC060228
DEFINITION	AC060228 Homo sapiens chromosome 4 clone RP11-922D2, WORKING DRAFT SEQUENCE.
ACCESSION	AC060228
VERSION	AC060228.10 GI:20335630
KEYWORDS	HTG; HTGS PHASE1; HTGS _DRAFT.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 157836) Wuzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
AUTHORS	

1 (bases 1 to 157836)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Alt-olsman, F.R., Allen, C., Albrooks, S.L., AmaralTunige, H.C., Are, V.R., Ayala, M., Banks, T., Barbieri, J., Benton, J., Bimsge, K., Blankenburg, K., Bonnan, D., Bouček, J., Bowles, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhey, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denu, A.L., Ding, Y., Dihn, H.J., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J.J., Garcia, A., Garner, T., Garza, N., Gill, R., Gortrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,  
 Homel, F., Howard, S., Huber, J., Huylk, S., Hume, J., Jackson, L.B.,  
 Jacobson, B., Jia, Y., Johnson, R., Jollivel, S., Jondah, S.,  
 Katsibon, B., Kelly, S., Khan, U., King, L., Kovari, J., Kovar, C.,  
 Kraybill, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loubes, H.,  
 Locado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maneshwar, M., Mapua, P., Martin, R., Martindale, A., Martinez, B.,  
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, N., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, N., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokhekwo, S., Ogun, M., Okunolu, G.,  
 Olegunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peeters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shoobart, N., Sison, I.,  
 Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,  
 Suton, A., Swatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Umanali, K., Vaquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wlaczek, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 157836)  
 Worley, K.C.  
 Direct Submission  
 Submitted (20-APR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 157836)  
 Worley, K.C.  
 Direct Submission  
 Submitted (08-JUN-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Apr 28, 2002 this sequence version replaced gi:16117945.  
 -----  
 Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 -----  
 Project Information  
 Center project name: HBBB  
 Center clone name: RP11-922D2  
 -----  
 Summary Statistics  
 Sequencing vector: M13:  
 Chemistry: Dye-primer Body: 93% of reads  
 Chemistry: Dye-terminator Big Dye: 7% of reads  
 Assembly program: Phrap; version 0.990139  
 Consensus quality: 147257 bases at least Q40  
 Consensus quality: 151545 bases at least Q40  
 Consensus quality: 153682 bases at least Q20  
 Estimated insert size: 166756; sum-of-contigs estimation  
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html))  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 29 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 2348: contig of 2348 bp in length  
 \* 2349 2448: gap of unknown length  
 \* 2449 4666: contig of 2218 bp in length  
 \* 4667 4766: gap of unknown length  
 \* 4767 6990: contig of 2224 bp in length

	*	6991	: gap of unknown length	
	*	7091	: contig of 2381 bp in length	
	*	9471	: gap of unknown length	
	*	9472	: gap of unknown length	
	*	9571	: contig of 2191 bp in length	
	*	11762	: gap of unknown length	
	*	11863	: contig of 3527 bp in length	
	*	15390	: gap of unknown length	
	*	15490	: contig of 3883 bp in length	
	*	19373	: gap of unknown length	
	*	19473	: contig of 3040 bp in length	
	*	22513	: gap of unknown length	
	*	22613	: contig of 2767 bp in length	
	*	25380	: gap of unknown length	
	*	25480	: contig of 4393 bp in length	
	*	29873	: gap of unknown length	
	*	29973	: contig of 4079 bp in length	
	*	34052	: gap of unknown length	
	*	34152	: contig of 3299 bp in length	
	*	37451	: gap of unknown length	
	*	37551	: contig of 4800 bp in length	
	*	42351	: gap of unknown length	
	*	42451	: contig of 4310 bp in length	
	*	46751	: gap of unknown length	
	*	46861	: contig of 3537 bp in length	
	*	50398	: gap of unknown length	
	*	50498	: contig of 2640 bp in length	
	*	53138	: gap of unknown length	
	*	53238	: contig of 3549 bp in length	
	*	56787	: gap of unknown length	
	*	56887	: contig of 4536 bp in length	
	*	61423	: gap of unknown length	
	*	61523	: contig of 4870 bp in length	
	*	66393	: gap of unknown length	
	*	66493	: contig of 6300 bp in length	
	*	72793	: gap of unknown length	
	*	72893	: contig of 6194 bp in length	
	*	79087	: gap of unknown length	
	*	79187	: contig of 7362 bp in length	
	*	86549	: gap of unknown length	
	*	86649	: contig of 4961 bp in length	
	*	91610	: gap of unknown length	
	*	91710	: contig of 8266 bp in length	
	*	99976	: gap of unknown length	
	*	100076	: contig of 8450 bp in length	
	*	108526	: gap of unknown length	
	*	108626	: contig of 5136 bp in length	
	*	113762	: gap of unknown length	
	*	113862	: contig of 12726 bp in length	
	*	126587	: gap of unknown length	
	*	126688	: contig of 14560 bp in length	
	*	141347	: gap of unknown length	
	*	141248	: contig of 16489 bp in length.	
	*	141348	: Location/Qualifiers	
		1..157836	/organism="Homo sapiens"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:9606"	
			/chromosome="4"	
			/clone="RP11-922D2"	
BASE COUNT	51525 a	26837 c	25711 g	50951 t 2812 others
ORIGIN				
Query Match	57.1%	Score 575.2;	DB 2;	Length 157836;
Best Local Similarity	99.5%;	Pred. No. 2.3e-106;		
Matches 577;	Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
OY	199	TGAGATGACATGACGATGGTCACCACCTTCAATCCATCTCGAATTATTCCTTAGGCAT	258	
DB	35791	TTTGAGATGACAAAGAGATGGTCACCCACTTATCATCTCGAATTATTCCTTAGGCAT	35850	
OY	259	ACGGAATTTTACCACTTCCTTTATTAATGCCCAGATGAAATACAGTCCCAAGTTACCTGG	318	
DB	35851	ACGGAATTTTACCACTTCCTTTATTAATGCCCAGATGAAATACAGTCCCAAGTTACCTGG	35910	

QY 319 GAATACCTTACACTGACACAGGGTTACCTTGTATCCCTGGATTCTTCTCTGAAAT 378  
 DB 35911 GAATACCTTACACTGACACAGGGTTACCTTGTATCCCTGGATTCTTCTCTGAAAT 35970  
 QY 379 CCCCTATGCTATCAATCCGTGGTTTCCCTTACCTACTCACTGTAATGTTCTCTCT 438  
 DB 35971 CCCCTATGCTATCAATCCGTGGTTTCCCTTACCTACTCACTGTAATGTTCTCTCT 36030  
 QY 439 CCCCTATGAGGGTTCCCGTTTGTCCCTCTCAAGGTTTTTTTGACAGCTGACAGACC 498  
 DB 36031 CCCCTATGAGGGTTCCCGTTTGTCCCTCTCAAGGTTTTTTTGACAGCTGACAGACC 36090  
 QY 499 CGCTGCCCCACCTATTTGACAGCTGAGCTGAGCTGACCTCTTACAGCCACACCTGT 558  
 DB 36091 CGCTGCCCCACCTATTTGACAGCTGAGCTGAGCTGACCTCTTACAGCCACACCTGT 36150  
 QY 559 AGAGACTGAGCTGCTGACAGGGGCCCTGTTGACAGCTGAGCTGAGCTGAGCTGCTGT 618  
 DB 36151 AGAGACTGAGCTGCTGACAGGGGCCCTGTTGACAGCTGAGCTGAGCTGAGCTGCTGT 36210  
 QY 619 TGAGAGCTGAGCTGCTGACAGGGCACCTGTGACAGCTGAGCTGAGGGACACCTGT 678  
 DB 36211 TGAGAGCTGAGCTGCTGACAGGGCACCTGTGACAGCTGAGCTGAGGGACACCTGT 36270  
 QY 679 TGAGAGTGAAGCCAGCTGACAGAGAACCTTACACAGCTGAGCTGCTACAGAGCCAGCTGC 738  
 DB 36271 TGAGAGTGAAGCCAGCTGACAGAGAACCTTACACAGCTGAGCTGCTACAGAGCCAGCTGC 36330  
 QY 739 TGCCCCAGAACTGACACCTCTTCTCCCTCTCTTGAACAGGCA 778  
 DB 36331 TGCCCCAGAACTGACACCTCTTCTCCCTCTCTTGAACAGGTA 36370

RESULT 6  
 BD109282 639 bp DNA linear PAT 18-SEP-2002  
 LOCUS BD109282  
 DEFINITION EST and encoded human protein.  
 ACCESSION BD109282.1 GI:23204100  
 VERSION BD109282.1  
 KEYWORDS JP 2002010789-A/1359.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 639)  
 Edwards, J.B.D.M., Jobert, S. and Giordano, J.B.  
 EST and encoded human protein  
 Patent: JP 2002010789-A 1359 15-JAN-2002;  
 GENSET CORP  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 OS Homo sapiens (human)  
 PN JP 2002010789-A/1359  
 PD 15-JAN-2002 JP 20020280989  
 PF 07-AUG-2000 JP 20020280989  
 PR 05-AUG-1999 US 60/147439  
 PI JEAN BAPTISTE DUMAS MIKNE EDWARDS, SEVEILIN JOBERT, JEAN EYE PI  
 GIORDANO  
 PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
 C12N1/21,  
 PC C12N5/10, C12P21/02, C12P21/08, C12O1/68, C12N15/00, C12N5/00, PC  
 C12N15/00  
 CC Von Heljne matrix  
 CC score 10.5  
 CC seq LLMACIVCVAFA/RK  
 FH Key Location/Qualifiers  
 FT CDS 185..613  
 FT 819 peptide 185..229.  
 FT Location/Qualifiers  
 1..639  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 158 a 164 c 118 g 198 t 1 others

## ORIGIN

Query Match 55.4%; Score 558; DB 6; Length 639;  
 Best Local Similarity 99.3%; Pred. No. 4, 1e-103;  
 Matches 580; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

QY 1 AGACAGACTAAAAAGCCATGATTCCTTCTCTCTTAAAGAAAGAAATATATAT 60  
 DB 58 AGACAGACTAAAAAGCCATGATTCCTTCTCTCTTAAAGAAAGAAATATATAT 117  
 QY 61 TAAATATATATGCTATTTTCTTAAACATTAATTAATGTTATATTTATAGAGTC 120  
 DB 118 TAAATATATGCTATTTTCTTAAACATTAATTAATGTTATATTTATAGAGTC 177  
 QY 121 AATCAAAATGAAGCTTCTCTTGGGCGCTGACATGATGTTGCTTTTGAAGAGAG 180  
 DB 178 AATCAAAATGAAGCTTCTCTTGGGCGCTGACATGATGTTGCTTTTGAAGAGAG 237  
 QY 181 ACGGTTCCCTTCATTTGAGATGACATAGAGATGACCACTTCATTCATCTCT 240  
 DB 238 ACGGTTCCCTTCATTTGAGATGACATAGAGATGACCACTTCATTCATCTCT 297  
 QY 241 GAATATTCCTTATGACATACGAAATTTACACCTCTCTTATATATGCGCCAGTAAAC 300  
 DB 298 GAATATTCCTTATGACATACGAAATTTACACCTCTCTTATATATGCGCCAGTAAAC 357  
 QY 301 AGTCCCGAGTTACCTGGAATATTAACACTGACACAGAGGTTACCTTCTATCCCTGAT 360  
 DB 358 AGTCCCGAGTTACCTGGAATATTAACACTGACACAGAGGTTACCTTCTATCCCTGAT 417  
 QY 361 TCTAACTTCTCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGAT 420  
 DB 418 TCTAACTTCTCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGAT 477  
 QY 421 GTTGAATGTTCT 480  
 DB 478 GTTGAATGTTCT 537  
 QY 481 TTCAGACAGCTGACAGACCGCTGACCCACCTATTTGACAGCTGAGCTGACAGCTGAC 540  
 DB 538 TTCAGACAGCTGACAGACCGCTGACCCACCTATTTGACAGCTGAGCTGACAGCTGAC 596  
 QY 541 TCTTACAGCCACACCTGTTAGACAGCTGAGCTGCTGACAGGGGCC 584  
 DB 597 TCTTACAGCCACACCTGTTAGACAGCTGAGCTGCTGACAGGGGCC 639

RESULT 7  
 AC069037 185969 bp DNA linear HTG 17-MAY-2000  
 LOCUS AC069037  
 DEFINITION Homo sapiens chromosome 4 clone RP11-653L5 map 4, WORKING DRAFT  
 SEQUENCE, 34 unordered pieces.  
 ACCESSION AC069037.1 GI:788485  
 VERSION AC069037.1  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 185969)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 4, clone RP11-653L5  
 Unpublished  
 2 (bases 1 to 185969)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, P.,  
 Boguslavskiy, L., Bouhassira, B., Brown, A., Burkett, G.,  
 Campopiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,  
 Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hagos, B., Haefford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karacas, A.,



TITLE  
JOURNAL

## COMMENT

Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lien, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,  
McCarthy, M., McEwan, P., McGarr, A., McKernan, K., McPheters, R.,  
Melartin, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, T., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Ollivar, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,  
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (17-MAY-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smith, A.P.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

## Project Information

Center project name: L7416

Center clone name: 653 L 5

## Summary Statistics

Sequencing vector: M13, M77815, 100% of reads

Chemistry: Dye-terminator Big Dye, 100% of reads

Assembly program: Phrap, version 0.960731

Consensus quality: 167163 bases at least Q40

Consensus quality: 175669 bases at least Q30

Consensus quality: 179699 bases at least Q20

Insert size: 18800; agarose-fp

Insert size: 182669; sum-of-contrigs

Quality coverage: 3.6 in Q20 bases; sum-of-contrigs

Quality coverage: 3.7 in Q20 bases; sum-of-contrigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 34 contrigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contrigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      1727: contrig of 1727 bp in length
*      1728      1827: gap of 100 bp
*      1728      3423: contrig of 1596 bp in length
*      3424      3523: gap of 100 bp
*      3524      6001: contrig of 2478 bp in length
*      6002      6101: gap of 100 bp
*      6102      7881: contrig of 1780 bp in length
*      7882      7981: gap of 100 bp
*      7982      10379: contrig of 2398 bp in length
*      10380      10479: gap of 100 bp
*      10480      13080: contrig of 2601 bp in length
*      13081      13180: gap of 100 bp
*      13181      15812: contrig of 2632 bp in length
*      15813      15912: gap of 100 bp
*      15913      18292: contrig of 2380 bp in length
*      18293      18392: gap of 100 bp
*      18393      20550: contrig of 2158 bp in length
*      20551      20650: gap of 100 bp
*      20651      24116: contrig of 3466 bp in length
*      24117      24216: gap of 100 bp
*      24217      27083: contrig of 2867 bp in length
*      27084      27183: gap of 100 bp
*      27184      31082: contrig of 3899 bp in length
*      31083      31182: gap of 100 bp
*      31183      36255: contrig of 5073 bp in length
*      36256      36355: gap of 100 bp
*      36356      40430: contrig of 4075 bp in length

```

## FEATURES

## source

```

40431      40530: gap of 100 bp
*      40531      44416: contrig of 3886 bp in length
*      44417      44516: gap of 100 bp
*      44517      48353: contrig of 3837 bp in length
*      48354      48453: gap of 100 bp
*      48454      52802: contrig of 4349 bp in length
*      52803      52902: gap of 100 bp
*      52903      58921: contrig of 6089 bp in length
*      58922      59091: gap of 100 bp
*      59092      63363: contrig of 4272 bp in length
*      63364      63463: gap of 100 bp
*      63464      68027: contrig of 4564 bp in length
*      68028      68127: gap of 100 bp
*      68128      72911: contrig of 4784 bp in length
*      72912      73011: gap of 100 bp
*      73012      79146: contrig of 6135 bp in length
*      79147      79246: gap of 100 bp
*      79247      86421: contrig of 7175 bp in length
*      86422      86521: gap of 100 bp
*      86522      92006: contrig of 5485 bp in length
*      92007      92106: gap of 100 bp
*      92107      97911: contrig of 5805 bp in length
*      97912      98011: gap of 100 bp
*      98012      103109: contrig of 5098 bp in length
*      103110      103209: gap of 100 bp
*      103210      109911: contrig of 6702 bp in length
*      109912      110011: gap of 100 bp
*      110012      119453: contrig of 9442 bp in length
*      119454      119553: gap of 100 bp
*      119554      126958: contrig of 7405 bp in length
*      126959      127058: gap of 100 bp
*      127059      134408: contrig of 7350 bp in length
*      134409      134508: gap of 100 bp
*      134509      143839: contrig of 9331 bp in length
*      143840      143939: gap of 100 bp
*      143940      153329: contrig of 11390 bp in length
*      153330      155429: gap of 100 bp
*      155430      168024: contrig of 12595 bp in length
*      168025      168124: gap of 100 bp
*      168125      185969: contrig of 17845 bp in length.

FEATURES
source
1..185969
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="4"
  /map="4"
  /clone="RP11-653L5"
  /clone_id="RP11 Human Male BAC"
1..1727
  /note="assembly_fragment"
1828..3423
  /note="assembly_fragment"
3524..6001
  /note="assembly_fragment"
6102..7881
  /note="assembly_fragment"
7982..10379
  /note="assembly_fragment"
10480..13080
  /note="assembly_fragment"
13181..15812
  /note="assembly_fragment"
15913..18292
  /note="assembly_fragment"
18393..20550
  /note="assembly_fragment"
20651..24116
  /note="assembly_fragment"
24217..27083
  /note="assembly_fragment"
27084..27183
  /note="assembly_fragment"
27184..31082
  /note="assembly_fragment"
31083..36255
  /note="assembly_fragment"
36256..40430
  /note="assembly_fragment"
vector_side:left"

```



```

misc_feature 27184..31082
            /note="assembly_fragment"
misc_feature 31183..36255
            /note="assembly_fragment"
misc_feature 36356..40430
            /note="assembly_fragment"
misc_feature 40531..44416
            /note="assembly_fragment"
misc_feature 44517..48353
            /note="assembly_fragment"
misc_feature 48454..52802
            /note="assembly_fragment"
misc_feature 52903..58991
            /note="assembly_fragment"
misc_feature 59092..63363
            /note="assembly_fragment"
misc_feature 63464..68027
            /note="assembly_fragment"
misc_feature 68128..72911
            /note="assembly_fragment"
misc_feature 73012..79146
            /note="assembly_fragment"
misc_feature 79247..86421
            /note="assembly_fragment"
misc_feature 86522..92006
            /note="assembly_fragment"
misc_feature 92107..97911
            /note="assembly_fragment"
misc_feature 98012..103109
            /note="assembly_fragment"

```

Query Match 53.3%; Score 537.6; DB 2; Length 185969;  
 Best Local Similarity 87.6%; Pred. No. 1e-98;  
 Matches 610; Conservative 0; Mismatches 84; Indels 2; Gaps 2;

```

QY 203 GATGACAATGACATGCTACCCACCTTCATCTCTGAAATTTCTTATGCAATACG 262
DB 167946 GATGACAATGACATGCTACCCACCTTCATCTCTGAAATTTCTTATGCAATACG 167887
QY 263 AATTTCACACCTCTCTTATTTATGCGCCAGATATACAGTCCAGTTACCTGGGAAT 322
DB 167886 AATTTCACACCTCTCTTATTTATGCGCCAGATATACAGTCCAGTTACCTGGGAAT 167827
QY 323 ACTTACACTGACACAGGGTTACCTTCGTATCCCTGATTTCTTAACTTCTCTGATTCGCC 382
DB 167826 ACTTACACTGACACAGGGTTACCTTCGTATCCCTGATTTCTTAACTTCTCTGATTCGCC 167767
QY 383 TATGTCTATCACTCCGTGTTTCCCTTACGTACAGTTGAATGTTCTCTCTCCCT 442
DB 167766 TATGTCTATCACTCCGTGTTTCCCTTACGTACAGTTGAATGTTCTCTCTCCCT 167707
QY 443 CCTAGGGGTTTCCGTTTCCCTTCCTCAAGTTTCTTTCAGCAGCTGACAGCCGCT 502
DB 167706 CCTA-GGGTTTCCGTTTCCCTTCCTTC-AGGTTTTCAGCAGCTGACAGCCGCT 167649
QY 503 GCCCACAATTGACAGCTGAGCTGCTGACAGCTGACCTCTTACAGCCACACTGTAGCA 562
DB 167648 GCCCACAATTGACAGCTGAGCTGCTGACAGCTGACCTCTTACAGCCACACTGTAGCA 167589
QY 563 GCTGAGCTGCTGACAGGGGCCCTGTTGACAGCTGAGCTGCTGACAGGACCTGTTGGA 622
DB 167588 GCTGAGCTGCTGACAGGGGCCCTGTTGACAGCTGAGCTGCTGACAGGACCTGTTGGA 167529
QY 623 GCTGAGCTGCTGACAGGACCTGTTGACAGCTGAGCTGCTGACAGGACCTGTTGGA 682
DB 167528 GCTGAGCTGCTGACAGGACCTGTTGACAGCTGAGCTGCTGACAGGACCTGTTGGA 167469
QY 683 GTGAGACCACTGCTGACAGAACTTCAACAGCTGACCTGTCTACAGCCAGCTGTGCTC 742
DB 167468 GTGAGACCACTGCTGACAGAACTTCAACAGCTGACCTGTCTACAGCCAGCTGTGCTC 167409
QY 743 CCGAATCTCACTCTCTCTCTCTTGAACAGGCAATGAGAAATCTCTAGAGAG 802
DB 167409 CCGAATCTCACTCTCTCTCTCTTGAACAGGCAATGAGAAATCTCTAGAGAG 802

```

```

DB 167408 CCGAATCTCACTCTCTCTCTCTTGAACAGGCAATGAGAAATCTCTAGAGAG 167349
QY 803 TACCATGGGTTCATTTCTTACTGACGAAATAGTAATCTCAAAAGTTTCTT 862
DB 167348 ATGTATGAGAAATGAGATTTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 167289
QY 863 CTTTTCAGAAAGCTATTTTCATTTCTGTTTATTCAGA 898
DB 167288 ATATCAACTTAATATTTTGAACAGCTTATTCAGA 167253

```

```

RESULT 8
BD107961
LOCUS BD107961 544 bp DNA linear PART 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD107961
VERSION BD107961.1 GI:23202779
KEYWORDS JP 2002010789-A/38.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 38 15-JAN-2002;
GENSET CORP

```

```

COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/38
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JUAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE PI
GIORDANO

```

```

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heljne matrix
CC score 10.5
CC seq LLMACIVCAFA/RX
FH Key Location/Qualifiers
FT CDS 185..544
FT sig_peptide 185..229.
FEATURES
source 1..544
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

BASE COUNT 142 a 127 c 93 g 182 t
ORIGIN

```

Query Match 46.6%; Score 470; DB 6; Length 544;  
 Best Local Similarity 99.8%; Pred. No. 3.1e-85;  
 Matches 481; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 1 AGACAGACTAAAGGAGGATGATCTTCTGTTCTCTTAAAGAGAGAAATATATT 60
DB 58 AGACAGACTAAAGGAGGATGATCTTCTGTTCTCTTAAAGAGAGAGAAATATATT 117
QY 61 TAAATAATCATTCGTAATTTCTTAAACATTAATTAATGTTTAATATGAGGTC 120
DB 118 TAAATAATCATTCGTAATTTCTTAAACATTAATTAATGTTTAATATGAGGTC 177
QY 121 AATCAAAATGAAGCTTCTCTTGGGCTGCAATTAATGTTTGTGTTTGAAGAGG 180
DB 178 AATCAAAATGAAGCTTCTCTTGGGCTGCAATTAATGTTTGTGTTTGAAGAGG 237
QY 181 ACGGTTCCCTTCATTTGTTGAGATGACATGACAGATGATGACCACTTCACTCT 240
DB 238 ACGGTTCCCTTCATTTGTTGAGATGACATGACAGATGATGACCACTTCACTCT 297
QY 241 GAAATTCCTTATGCAATGCAATTTTACCACTCTCTTATTTATGCCCAAGTAATAC 300

```

[illegible]

Dd		421	GCAACGCGGTGGAGCTGAAGCCCTGCAGAGGACACTTGTTCAGCTGAAGCTTGCTGCA	667
Oy		608	GAGGACCCTGTGGAGCTGAAGCCCTGCAGAGGACACTTGTTCAGCTGAAGCTTGCTGCA	667
Dd		481	GARGCNCNGTGNANGNCGARCCNGCNGGARCCNCGNTGNCNGNCGARCCNGCNGCN	540
Oy		668	GAGCACCTGTGGAGCTGAAGCCCTGCAGAGGACACTTGTTCAGCTGAAGCTTGCTGCA	727
Dd		541	GARGCNCNGTGNANGNCGARCCNGCNGGARCCNCGNTGNCNGNCGARCCNGCNGCN	600
Oy		728	GCCAAGCTGTGCCCGGAGAAGCTGCAAGCTTTCTCTCTTTGAAACAGCAATCAG	784
Dd		601	GCNAARCNGNGCNCNGARCCNCAICCMNSNCMNSNTTNGARCARGCNAATCAR	657
RESULT 10				
LOCUS	BD071746	470 bp	DNA	linear PAT 27-AUG-2002
DEFINITION	Secreted expressed sequence tags (ESTs).			
ACCESSION	BD071746			
VERSION	BD071746.1 GI:22617349			
KEYWORDS	JP 2001519667-A/556.			
SOURCE	unidentified			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 470)			
AUTHORS	Jacobs,K., McCoy,J.M., Lavallie,B.R., Racie,L.A., Merberg,D., Treacy,M., Spaulding,V., and Agostino,M.J.			
TITLE	Secreted expressed sequence tags (ESTs)			
JOURNAL	Patent: JP 2001519667-A 556 23-OCT-2001; GENETICS INSTITUTE INC			
COMMENT	OS Unidentified PN JP 2001519667-A/556 PD 23-OCT-2001 PF 10-APR-1998 JP 1998543069 PR 10-APR-1997 US 08/838821 PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI DAVID MERBERG, PI MAURICE TREACY,VIKIRI SPAULDING,MICHAEL J AGOSTINO PC C12N5/12,C12N5/10,C07K14/47,C12Q1/68,A6IX38/17 CC Strandedness: Double; CC Topology: linear; CC Secreted expressed sequence tags (ESTs) FH Key Location/Qualifiers FT source 1..470 FT /organism='Unidentified'.			
FEATURES				
source	location/qualifiers 1..470 /organism="unidentified" /mol_type="genomic DNA"			
BASE COUNT	116 a 115 c 83 g 156 t			
ORIGIN				
Query Match	44.5%; Score 449; DB 6; Length 470;			
Best Local Similarity	98.9% Pred. No. 5.7e-81;			
Matches	452; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
Oy	11 AAAAAGCATGATTCTTTCGTTCTCTCTAAAAGAAAAATAATTAATTAATAATAACA	70		
Dd	11 AAAGAGCCTGATATCTTCCTCTCTCTAAAAGAAAAATAATTAATTAATAATAACA	70		
Oy	71 TTGGATATTTCTAAAACATAAATTAATTAATGTATTAATTCATAGGTCATCAAATG	130		
Dd	71 TTGGATATTTCTAAAACATAAATTAATTAATGTATTAATTCATAGGTCATCAAATG	130		
Oy	131 AAGCTTCTCCCTTAGGCTGCATGTATGTGTTCTTTTGCAAGAAAGACGGTTCCCC	190		
Dd	131 AAGCTTCTCCCTTAGGCTGCATGTATGTGTTCTTTTGCAAGAAAGACGGTTCCCC	190		
Oy	191 TTCATTTGTGAGATGACATGACATGATGATCCCACCTTCATTCATCTTGAATATTCT	250		

Db	191	TTCAATGGTAGGATGACAAATGAGATGTCACCCACTTCATTCATCTCTGAATATTCCT	250
Qy	251	TATGGCATAGCGGAATTTCCACCTCTCTTTATTTATGCGCCAGTAATAGAGTCCCCAGT	310
Db	251	TATGGCAATACGGAATTTACCACTCTCTTTATTTATCCCAAGTAATAGAGTCCCCAGT	310
Qy	311	TACCTGGGAATACCTTACACTGACACAGAGGTACCTTGTAATCCCTGGATTTCTAATCTCT	370
Db	311	TACCTGGGAATACCTTACACTGACACAGAGGTACCTTGTAATCCCTGGATTTCTAATCTCT	370
Qy	371	CCTGGATTTCCCTATGTCATACATCCGATGTTTTCCCTTAGCTACTACATTAATGTT	430
Db	371	CCTGGATTTCCCTATGTCATACATCCGATGTTTTCCCTTAGCTACTACATTAATGTT	430
Qy	431	CCTCTCTCTCCCTCCTAGAGGGTTTTCCCGTTTGTCCTC	467
Db	431	CCTCTCTCTCCCTCCTAGAGGGTTTTCCCGTTTGTCCTC	467
RESULT 11			
LOCUS	BD071795	450 bp	DNA
DEFINITION	Secreted expressed sequence tags (sESTs).	linear	PAT 27-AUG-2002
ACCESSION	BD071795		
VERSION	BD071795.1	GI:22617398	
KEYWORDS	JP 2001519667-A/605.		
SOURCE	unidentified		
ORGANISM	unidentified.		
REFERENCE	1 (bases 1 to 450)		
AUTHORS	Jacobs, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,		
TITLE	Tracy, M., Spaulding, V. and Apostino, M.J.		
JOURNAL	Secreted expressed sequence tags (sESTs)		
COMMENT	Patent: JP 2001519667-A 605 23-OCT-2001;		
	GENETICS INSTITUTE INC		
	OS Unidentified		
	PN JP 2001519667-A/605		
	PD 23-OCT-2001		
	PF 10-APR-1998 JP 1998543069		
	PR 10-APR-1997 US 08/838821		
	PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI		
	DAVID MERBERG,		
	PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC		
	CIN15/12, CIN15/10, C07K14/47, C12Q1/68, A61K38/17 CC		Strandedness:
	Double;		
	CC Topology: Linear;		
	CC Secreted expressed sequence tags (sESTs)		
	FH Key Location/Qualifiers		
	FT source 1.450		
	FT Location/Qualifiers		
	1.450		
	/organism="unidentified"		
	/source="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
	BASE COUNT 112 a 113 c 78 g 147 t		
	ORIGIN		
	Query Match 41.8%; Score 421; DB 6; Length 450;		
	Query Local Similarity 100.0%; Pred. NO. 2.8e-75; Indels 0; Gaps 0;		
	Matches 421; Conservative 0; Mismatches 0;		
Qy	49	AAAAATATATTTAAAAATACATGCGTATTTTCTAAAAACAATAATTAATAGTGTAT	108
Db	21	AAAAATATATTTAAAAATACATGCGTATTTTCTAAAAACAATAATTAATAGTGTAT	80
Qy	109	ATTCAATAGGTCATCAAAATGAGAGCTTCTCTTTGGGCTGCAATGTAATGTTGCTTT	168
Db	81	ATTCAATAGGTCATCAAAATGAGAGCTTCTCTTTGGGCTGCAATGTAATGTTGCTTT	140
Qy	169	TGCAAGAGAGAGAGCGTTCCCTTCATTTGTTGAGGATGACAAATGACGATGTCACCACT	228
Db	141	TGCAAGAGAGAGAGCGTTCCCTTCATTTGTTGAGGATGACAAATGACGATGTCACCACT	200

QY		229	TGATCATCTGTAAATTTCCTTAAGCACAAGGAATTACCAACCCTCTTATTAATGC	288
Db		201	TGATCATCTGTAAATTTCCTTAAGCACAAGGAATTACCAACCCTCTTATTAATGC	260
QY		289	CCGAGTGAAATACAGTCCCAGTTACCTGGGAACTACTTACATGACACAGAGTTACCTTC	348
Db		261	CCGAGTGAAATACAGTCCCAGTTACCTGGGAACTACTTACATGACACAGAGTTACCTTC	320
QY		349	GTAATCCTGGAATTTCACTTCTCTCGAATTCCTCTAATGTCTATCACATCCGTGTTTTCC	408
Db		321	GTAATCCTGGAATTTCACTTCTCTCGAATTCCTCTAATGTCTATCACATCCGTGTTTTCC	380
QY		409	CTTAGCTACTCAGTTGAATGTTCCTCTCCCTCCCTAGGAGGTTTCCGTTTGCCTTC	468
Db		381	CTTAGCTACTCAGTTGAATGTTCCTCTCCCTCCCTAGGAGGTTTCCGTTTGCCTTC	440
QY		469	T 469	
Db		441	T 441	
RESULT 12				
AC069037			185969 bp DNA linear HTG 17-MAY-2000	
LOCUS			Homo sapiens chromosome 4 clone RP11-653L5 map 4, WORKING DRAFT	
DEFINITION			SEQUENCE, 34 uncloned pieces.	
ACCESSION			AC069037 GI:7884885	
VERSION			HTG; HTG_PHAISI; HTGS_DRAFT.	
KEYWORDS			Homo sapiens (human)	
SOURCE			Homo sapiens	
ORGANISM			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE			1 (bases 1 to 185969)	
AUTHORS			Birren,B., Linton,L., Nuebaum,C. and Lande,E.	
JOURNAL			Homo sapiens chromosome 4, clone RP11-653L5	
REFERENCE			Unpublished	
AUTHORS			2 (bases 1 to 185969)	
			Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barra,N., Baetien,V., Bedd,F., Boguslavsky,I., Boukhalter,B., Brown,A., Burkett,G., Campioano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,D.S., Dodge,S., Domino,M., Doyle,W., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Glnde,S., Goylete,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Hottom,L., Howland,J.C., Illey,I., Johnson,R., Jones,C., Kann,L., Karasab,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., McKernan,K., McGuire,N., McCarthy,M., McGwan,P., McGurk,K., McKernan,K., McPheeters,R., Moldrim,J., Menues,L., Mihova,T., Miranda,C., Miengs,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanik,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testfeyre,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zaitoun,J., Zimmer.A. and Zoody,M.	
TITLE			Direct Submission	
JOURNAL			Submitted (17-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA	
COMMENT			All repeats were identified using RepeatMasker: Smit, A.P.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a>	
			----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: <a href="http://www-seq.wi.mit.edu">http://www-seq.wi.mit.edu</a> Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: U7416	

```

Center clone name: 653 L5
----- Summary Statistics -----
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 167163 bases at least Q40
Consensus quality: 175669 bases at least Q30
Consensus quality: 179699 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 182669; sum-of-contigs
Quality coverage: 3.6 in Q20 bases; agarose-fp
Quality coverage: 3.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
1727: contig of 1727 bp in length
1728      1827: gap of 100 bp
1828      3423: contig of 1596 bp in length
3424      3523: gap of 100 bp
3524      6001: contig of 2478 bp in length
6002      6101: gap of 100 bp
6102      7881: contig of 1780 bp in length
7882      7981: gap of 100 bp
7982      10379: contig of 2398 bp in length
10380      10479: gap of 100 bp
10480      13080: contig of 2601 bp in length
13081      13180: gap of 100 bp
13181      15812: contig of 2632 bp in length
15813      15912: gap of 100 bp
15912      18292: contig of 2380 bp in length
18293      18392: gap of 100 bp
18392      20550: contig of 2158 bp in length
20551      20650: gap of 100 bp
20651      24116: contig of 3466 bp in length
24117      24216: gap of 100 bp
24217      27083: contig of 2867 bp in length
27084      27183: gap of 100 bp
27183      31082: contig of 3899 bp in length
31083      31182: gap of 100 bp
31182      36255: contig of 5073 bp in length
36256      36355: gap of 100 bp
36355      40430: contig of 4075 bp in length
40431      40530: gap of 100 bp
40531      44416: contig of 3886 bp in length
44417      44516: gap of 100 bp
44517      48353: contig of 3837 bp in length
48354      48453: gap of 100 bp
48453      52802: contig of 4349 bp in length
52803      52902: gap of 100 bp
52902      58991: contig of 6099 bp in length
58992      59091: gap of 100 bp
59091      63363: contig of 4272 bp in length
63364      63463: gap of 100 bp
63463      68027: contig of 4564 bp in length
68028      68127: gap of 100 bp
68127      72911: contig of 4784 bp in length
72912      73011: gap of 100 bp
73011      79146: contig of 6135 bp in length
79147      79246: gap of 100 bp
79247      86421: contig of 7175 bp in length
86422      86521: gap of 100 bp
86521      92006: contig of 5485 bp in length
92007      92106: gap of 100 bp
92107      97911: contig of 5805 bp in length
97912      98011: gap of 100 bp
98012      103109: contig of 5098 bp in length
103110      103209: gap of 100 bp

```

```

FEATURES
source
* 103210 109911: contig of 6702 bp in length
* 109912 110011: gap of 100 bp
* 110012 119453: contig of 9442 bp in length
* 119454 119553: gap of 100 bp
* 119554 126958: contig of 7405 bp in length
* 126959 127058: gap of 100 bp
* 127059 134408: contig of 7350 bp in length
* 134409 134508: gap of 100 bp
* 134509 143839: contig of 9331 bp in length
* 143840 143939: gap of 100 bp
* 143940 155329: contig of 11390 bp in length
* 155330 155429: gap of 100 bp
* 155430 168024: contig of 12595 bp in length
* 168025 168124: gap of 100 bp
* 168125 185965: contig of 17845 bp in length.
Location/Qualifiers
1. 185969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-653L5"
/clone_1bp="RP11 Human Male BAC"
1. 1727
/note="assembly_fragment"
1828. 3423
/note="assembly_fragment"
3524. 6001
/note="assembly_fragment"
6102. 7881
/note="assembly_fragment"
7982. 10379
/note="assembly_fragment"
10480. 13080
/note="assembly_fragment"
13181. 15812
/note="assembly_fragment"
15913. 18292
/note="assembly_fragment"
18393. 20550
/note="assembly_fragment"
20651. 24116
/note="assembly_fragment"
24217. 27083
/note="assembly_fragment"
clone_end:17
vector_side:left"
27184. 31082
/note="assembly_fragment"
31183. 36255
/note="assembly_fragment"
36356. 40430
/note="assembly_fragment"
40531. 44416
/note="assembly_fragment"
44517. 48353
/note="assembly_fragment"
48454. 52802
/note="assembly_fragment"
52903. 58991
/note="assembly_fragment"
59092. 63363
/note="assembly_fragment"
63464. 68027
/note="assembly_fragment"
68128. 72911
/note="assembly_fragment"
73012. 79146
/note="assembly_fragment"
79247. 86421
/note="assembly_fragment"
86522. 92006
/note="assembly_fragment"

```

misc\_feature /note="assembly\_fragment"  
92107..97911  
/note="assembly\_fragment"  
misc\_feature /note="assembly\_fragment"  
98012..103109

Query Match 41.6%; Score 419.4; DB 2; Length 185969;  
Best Local Similarity 99.8%; Pred. No. 1.1e-74;  
Matches 420; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 TGAGATGAGAAAGACAGATGTCACCACTTCATCCATCTCTGAATATTCCTATATGGCAT 258  
DB 67607 TTAGAGAGACAAAGACATGATGTCACCACTTCATCCATCTCTGAATATTCCTATATGGCAT 67666  
QY 259 ACCGAATTTACCACTCTCTTATTATGSCCAAGTAATACAGTCCCAAGTACCCTGG 318  
DB 67667 ACCGAATTTACCACTCTCTTATTATGSCCAAGTAATACAGTCCCAAGTACCCTGG 67726  
QY 319 GAATACTTACATGACACAGGGTTACCTTCGTATCCCTGGATTCTTAACCTTCCTGGATT 378  
DB 67727 GAATACTTACATGACACAGGGTTACCTTCGTATCCCTGGATTCTTAACCTTCCTGGATT 67786  
QY 379 CCCCTATGCTATACATCCGATGTTTCCCTAGACTACAGTGAATGTTCTCTCT 438  
DB 67787 CCCCTATGCTATACATCCGATGTTTCCCTAGACTACAGTGAATGTTCTCTCTCT 67846  
QY 439 CCCCTAGAGGGTTCCGATGTTTCCCTCAAGGTTTTCAGACAGCTGACACACC 498  
DB 67847 CCCCTAGAGGGTTCCGATGTTTCCCTCAAGGTTTTCAGACAGCTGACACACC 67906  
QY 499 CGCTGCCACCTATGACAGCTGACCTGCTGACAGCTGACCTCTTACAGCCACACCTGT 558  
DB 67907 CGCTGCCACCTATGACAGCTGACCTGCTGACAGCTGACCTCTTACAGCCACACCTGT 67966  
QY 559 AGCAGCTGAGCCCTGCGACAGGGGCCCTGTTGACAGCTGAGCCGTCGACAGGACACCTGT 618  
DB 67967 AGCAGCTGAGCCCTGCGACAGGGGCCCTGTTGACAGCTGAGCCGTCGACAGGACACCTGT 68026  
QY 619 T 619  
DB 68027 T 68027

RESULT 13  
LOCUS BD122986 485 bp DNA linear PAT 18-SEP-2002  
DEFINITION EST and encoded human protein.  
ACCESSION BD122986  
VERSION BD122986.1 GI:23217931  
KEYWORDS JP 2002010789-A/15063.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 485)  
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.R.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 15063 15-JAN-2002;  
GENSET CORP  
OS Homo sapiens (human)  
PN JP 2002010789-A/15063  
PD 15-JAN-2002  
PR 07-AUG-2000 JP 2002280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT source 1..485

FEATURES  
source Location/Qualifiers  
1..485  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 126 a 121 c 92 g 144 t 2 others

Query Match 33.7%; Score 339.6; DB 6; Length 485;  
Best Local Similarity 96.4%; Pred. No. 9.7e-59;  
Matches 345; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 170 GCAAGAGAGAGACGGTTCCTTCATTTGGTGAAGTACATGACAGATGTCACCCACTT 229  
DB 128 GGAAGAGAGAGACGGTTCCTTCATTTGGTGAAGTACATGACAGATGTCACCCACTT 187  
QY 230 CATCCATCTCTGAATATTCCTATGACATAGCGAATTTACACCTCTCTTATATTCG 289  
DB 188 CATCCATCTCTGAATATTCCTATGACATAGCGAATTTACACCTCTCTTATATTCG 247  
QY 290 CCAGTGAATACAGTCCCAAGTACCCTGGAAATCTTACATGACACAGGGTTACCTCG 349  
DB 248 CCAGTGAATACAGTCCCAAGTACCCTGGAAATCTTACATGACACAGGGTTACCTCG 307  
QY 350 TATCCCTGATTTCTAATCTTCCTGATTCCTCTATGTCATACATCCGTGTTTCCC 409  
DB 308 TATCCCTGATTTCTAATCTTCCTGATTCCTCTATGTCATACATCCGTGTTTCCC 367  
QY 410 TTAGCTACTCAATGATATTCCTCTCTCCCTCTCAGAGGGTTTCCCGTTTCCCTCCT 469  
DB 368 TTAGCTACTCAATGATATTCCTCTCTCCCTCTCAGAGGGTTTCCCGTTTCCCTCCT 427  
QY 470 TCAAGGTTTTTTCAGACAGCTGACAGACCCGCTGCCCACTATTGACAGCTGACCTG 527  
DB 428 TCAAGGTTTTTTCAGACAGCTGACAGACCCGCTGCCCACTATTGACAGCTGACCTG 485

RESULT 14  
LOCUS BD071681 296 bp DNA linear PAT 27-AUG-2002  
DEFINITION Secreted expressed sequence tags (ESTs).  
ACCESSION BD071681  
VERSION BD071681.1 GI:22617284  
KEYWORDS JP 2001519667-A/491.  
SOURCE unidentified  
ORGANISM unidentified

REFERENCE 1 (bases 1 to 296)  
AUTHORS Jacobs, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Werberg, D.,  
Trecay, M., Spaulding, V. and Agostino, M.J.  
TITLE Secreted expressed sequence tags (ESTs)  
JOURNAL Patent: JP 2001519667-A 491 23-OCT-2001;  
GENETICS INSTITUTE INC  
OS Unidentified  
PN JP 2001519667-A/491  
PD 23-OCT-2001  
PR 10-APR-1998 JP 1998543069  
PR 10-APR-1997 US 08/838821  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
DAVID MERBERG,  
PI MAURICE TRECAY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: linear;  
CC Secreted expressed sequence tags (ESTs)  
FH Key Location/Qualifiers  
FT source 1..296  
FT Location/Qualifiers  
FT /organism="Unidentified".

FEATURES  
source Location/Qualifiers  
1..296  
/organism="unidentified"  
/mol\_type="genomic DNA"

```

BASE COUNT      50 a 99 c 82 g 65 t
ORIGIN
Query Match      28.4%; Score 286.4; DB 6; Length 296;
Best Local Similarity 99.7%; Prid. No. 66-48;
Matches 287; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 449 GGTTCCTCCCTTCCCTCTTCAAGTTTTCAGACGACGACCCCGTGCCTCA 508
DB 1 GGTTCCTCCCTTCCCTCTTCAAGTTTTCAGACGACGACCCCGTGCCTCA 60
QY 509 CCTATTGACGACGACGCTGCTGACGACGCTCTTACAGCCACACCTGTAGACGCTGAG 568
DB 61 CCTATTGACGACGACGCTGCTGACGACGCTCTTACAGCCACACCTGTAGACGCTGAG 120
QY 569 CCGTCGTCGACGAGGCGCCCTGTTGACGACGCTGCTGACGACGACCTGTTGAGCTGAG 628
DB 121 CCGTCGTCGACGAGGCGCCCTGTTGACGACGCTGCTGACGACGACCTGTTGAGCTGAG 180
QY 629 CCTGCTGACGACGACGCTGTTGACGACGCTGCTGACGACGACCTGTTGAGCTGAG 688
DB 181 CCGTCGTCGACGACGACGCTGTTGACGACGCTGCTGACGACGACCTGTTGAGCTGAG 240
QY 689 CCAAGCTGACGACGACGACCTTACAGCTGACGCTGCTTACAGCCAGGCT 736
DB 241 CCAAGCTGACGACGACGACCTTACAGCTGACGCTGCTTACAGCCAGGCT 288

RESULT 15
AC099443/c 259038 bp DNA 1linear HTG 10-MAY-2003
LOCUS Rattus norvegicus clone CH230-197D1, *** SEQUENCING IN PROGRESS
DEFINITION
AC099443
AC099443.5 GI:30521301
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 259038)
Muzny,D.,Marie., Metzker,W.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,S., Bacc,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,B.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregreggs,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,U.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Huilyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolyvet,A.,
Karachy,S., Kelly,S., Kelly,S., Khan,Z., King,J., Kovar,C.,
Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshues,L., Louisedge,H., Lozano,R.U., Lu,X., Ma,J.,
Manshewari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawlawy,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,

```

```

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
TITLE
SUBMITTED (15-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 259038)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:2364454.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHL
Center clone name: CH230-197D1
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 237512 bases at least Q40
Consensus quality: 237879 bases at least Q30
Consensus quality: 240733 bases at least Q20
Estimated insert size: 249015; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 244738: contig of 244738 bp in length

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 05:51:13 : Search time 336 seconds  
(without alignments)  
8098.314 Million cell updates/sec

Title: US-09-923-236-1

Perfect score: 1008  
Sequence: 1 agacagacacaaaagcat.....tagatatttagacacatg 1008

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756/seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

N\_Geneseq\_19jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1008	100.0	1008	24	AAD45050 Human zslg63 DNA.
2	1008	100.0	1008	24	ABBS2633 DNA encoding human
3	1008	100.0	1008	24	AA820591 Human zslg63 cDNA.
4	1008	100.0	1008	25	ABX93594 Human cDNA encodin
5	997	98.9	998	22	AA272729 Human transport pr
6	981.6	97.4	1325	20	AAV80740 Human secreted pro
7	896.4	88.9	959	20	AAK61352 DNA encoding a hum
8	460	45.6	657	24	AAD45051 Human zslg63 degen

9	460	45.6	657	24	ABBS2634
10	460	45.6	657	24	AA820592
11	460	45.6	657	25	ABX93595
12	449	44.5	470	20	AAV89597
13	421	41.8	450	20	AAV89646
14	286.4	28.4	296	20	AAV89532
15	121.2	12.0	3489	21	AAAJ0290
16	121.2	12.0	3489	22	AAAF8901
17	121.2	12.0	3489	24	AA893487
18	121.2	12.0	32807	20	AAV73805
19	121.2	12.0	137507	19	AAV19941
20	106	10.5	5120	22	AAAC8477
21	100	9.9	1037	21	AAAS9242
22	100	9.9	1472	21	AAAS9241
23	99.2	9.8	1159	21	AAAS9240
24	98.2	9.7	2108	24	ABL67774
25	93.6	9.3	397	20	AAK89891
26	92.8	9.2	5120	22	AAAC8467
27	92.4	9.2	1101	22	AAAD06303
28	92	9.1	543	13	AAQ23092
29	89.8	8.9	877	24	ABZ1932
30	89.8	8.9	3042	23	ABL28445
31	89.8	8.9	5215	23	ABL28444
32	88.4	8.8	403	21	ABK10240
33	88.2	8.8	627	23	AA893555
34	87.6	8.7	234	16	AAO84832
35	87.6	8.7	1101	22	AAAD06303
36	86.6	8.6	786	17	AAAT46148
37	86.6	8.6	786	20	AAZ10978
38	86.4	8.6	486	22	AAV75507
39	85.2	8.5	799	21	AAV55831
40	85.2	8.5	1926	21	AAAS0254
41	85.2	8.5	1926	22	AAAF82902
42	85.2	8.5	2580	21	AAAT75454
43	85.2	8.5	2580	24	AAI64275
44	85.2	8.5	5452	20	AAK90923
45	85.2	8.5	8705	20	AAZ233778

#### ALIGNMENTS

RESULT 1	AAAD5050	AAAD5050 standard; DNA; 1008 BP.
ID	AAAD5050	
XX	AAAD5050;	
AC		
XX		
DT	27-DEC-2002 (first entry)	
XX		
DE	Human zslg63 DNA.	
XX		
KW	Human; secreted salivary protein; zslg63 protein; host defense protein;	
KW	immune modulating factor; antipathogenic; cell-cell signalling molecule;	
KW	growth factor; cytokine; growth factor hormone activity; dental carrier;	
KW	infection; tooth decay; periodontal disease; gastrointestinal disease;	
KW	churn; urinary tract infection; vaginal infection; diabetes; obesity;	
KW	anti-inflammatory; chronic tissue damage; lung dysfunction; restenosis;	
KW	gene therapy; salivary gland dysfunction; prostate gland dysfunction;	
KW	forensic DNA profiling; chondrosarcoma; atherosclerosis; chromosome 4;	
KW	gene; de.	
XX		
OS	Homo sapiens.	
XX		
PH		
FT	Key	Location/Qualifiers
FT	CDS	128..787
FT		/*tag= a
FT		/product= "Human zslg63 protein"
FT	sig_peptide	128..172
FT		/*tag= b
FT	mat_peptide	173..784
FT		/*tag= c
FT		/product= "Mature human zslg63 protein"

Human secreted sal  
Human zslg63 degen  
Human zslg63 degen  
EST clone CP251.  
EST clone CP116.  
Kaposi's sarcoma-a  
Nucleotide sequenc  
Kaposi's sarcoma-a  
KSHV LTR DNA (nucl  
KSHV long unique c  
DNA sequence of hu  
Exons B, C and A o  
An EcoRI fragment  
Oesophagus cancer  
Spinocerebellar at  
DNA sequence of hu  
Human cDNA clone H  
Antigen tc-7a gene  
Human polymucleoti  
Drosophila melanog  
ABL28444  
Tritonucleotide repe  
DNA encoding novel  
Spinocerebellar at  
Human cDNA clone H  
T. cruzi L19E homo  
TCR coding sequenc  
Polyglutamine tranc  
Nucleotide sequenc  
Epstein Barr virus  
EBV tethering prot  
Nucleotide sequenc  
Epstein-Barr virus  
Anti-sense strand  
Vector pShuttle DN

XX US2002090677-A1.  
 XX 11-JUL-2002.  
 XX 03-AUG-2001; 2001US-0923236.  
 XX 17-MAR-1999; 99US-124820P.  
 XX 17-MAR-2000; 2000US-0527345.  
 XX  
 XX (ADIE/) ADLER D A.  
 XX (SHEP/) SHEPPARD P O.  
 XX  
 XX Adler DA, Sheppard PO;  
 XX  
 XX MPI: 2002-642378/69.  
 XX P-PDB; AAE27862.  
 XX  
 XX Novel secreted salivary polypeptide, zslg63, useful as antimicrobial  
 XX agent for treating microbial infection, dental caries, periodontal  
 XX disease, thrush gastrointestinal disease, and for aiding digestion -  
 XX  
 XX Claim 3, Page 27-28; 33pp; English.

CC The invention relates to human secreted salivary polypeptide designated  
 CC as zslg63 and nucleic acid molecules encoding such polypeptides. zslg63  
 CC can be used in detecting agonists and antagonists of its activity, and  
 CC is also useful as a host defense polypeptide, immune modulating factor,  
 CC antipathogenic polypeptide, cell-cell signalling molecule, growth factor,  
 CC cytokine, or as secreted extracellular matrix associated proteins with  
 CC growth factor hormone activity. It is useful for treating conditions  
 CC associated with pathological microbes, including bacterial, fungal and  
 CC viral infections, for treating dental caries (tooth decay), periodontal  
 CC disease, thrush and gastrointestinal disease, for treating urinary tract  
 CC infection, vaginal infection and for preventing infection in skin and  
 CC other epithelial wounds. zslg63 is useful for establishing normal  
 CC microflora and protect against pathogenic colonisation and invasion, for  
 CC treating chronic tissue damage e.g. damage in extremities associated  
 CC with diabetes and useful as anti-inflammatory agents. It is useful as a  
 CC marker of lung dysfunction, salivary gland dysfunction, or dysfunction of  
 CC prostate gland. It is also therapeutically useful for aiding digestion.  
 CC Polynucleotides of the invention are used in gene therapy for increasing  
 CC or inhibiting zslg63 activity, for detecting abnormalities on human  
 CC chromosome 4 associated with disease or other human traits and as  
 CC diagnostic in forensic DNA profiling. Sequences of the invention are  
 CC useful for stimulating proliferation or differentiation of cardiac  
 CC myocytes, for proliferation or differentiation of adipocytes and for  
 CC inhibiting chondrosarcomas, atherosclerosis, restenosis and obesity.  
 CC The present sequence is human zslg63 DNA. zslg63 gene is located on  
 CC chromosome 4.  
 CC  
 CC Sequence 1008 BP; 257 A; 251 C; 193 G; 307 T; 0 other;

Query Match 100.0%; Score 1008; DB 24; Length 1008;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-207;  
 Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACGAGCTAAAGGACGATGATCTTCTGTTCTCTAAAGAGAAATATTAAT 60  
 DB 1 AAGCAGACTTAAGGACGATGATCTTCTGTTCTCTAAAGAGAAATATTAAT 60  
 QY 61 TAAATAATCATTCGATTTTCTTAAACATTAATTAATAGTTAATATTCATAGGTC 120  
 DB 61 TAAATAATCATTCGATTTTCTTAAACATTAATTAATAGTTAATATTCATAGGTC 120  
 QY 121 AATCAAAATGAAGCTTCTCTTGGGCGCTGATGATGTGTTGCTTTGCAAGAAAG 180  
 DB 121 AATCAAAATGAAGCTTCTCTTGGGCGCTGATGATGTGTTGCTTTGCAAGAAAG 180  
 QY 181 ACGGTTCCCTTCATTTGATGAGATGACATGACGATGGTACCCATTCATCCATCTCT 240  
 DB 181 ACGGTTCCCTTCATTTGATGAGATGACATGACGATGGTACCCATTCATCCATCTCT 240

QY 241 GAATATTCCTTATGAGCATACGAAATTTACACCTCTCTTATTAATGCGCCAGTAAATAC 300  
 DB 241 GAATATTCCTTATGAGCATACGAAATTTACACCTCTCTTATTAATGCGCCAGTAAATAC 300  
 QY 301 AGTCCCAATTAACCTGGGAAATATTACACTGACACAGGGTTACTTGTATCCCTGGAT 360  
 DB 301 AGTCCCAATTAACCTGGGAAATATTACACTGACACAGGGTTACTTGTATCCCTGGAT 360  
 QY 361 TCTAATCTCTCGGAAATCCCTAATGCTATACATCCGAGGTTTCCCTAGCTACTCA 420  
 DB 361 TCTAATCTCTCGGAAATCCCTAATGCTATACATCCGAGGTTTCCCTAGCTACTCA 420  
 QY 421 GTTAAATGTTCTCTCTCTCCCTCCAGAGGTTTCCCTGTTCCCTCCCTCAAGTTT 480  
 DB 421 GTTAAATGTTCTCTCTCTCCCTCCAGAGGTTTCCCTGTTCCCTCCCTCAAGTTT 480  
 QY 481 TTCAAGAGCTGACAGACCCGCTGCCCACTTATTCAGAGCTGAGCTGTGACGTGAC 540  
 DB 481 TTCAAGAGCTGACAGACCCGCTGCCCACTTATTCAGAGCTGAGCTGTGACGTGAC 540  
 QY 541 TCTTACAGCCACACCTGTAGAGCTGAGGCTGTCAGAGGAGGAGGAGGAGGAGGAGG 600  
 DB 541 TCTTACAGCCACACCTGTAGAGCTGAGGCTGTCAGAGGAGGAGGAGGAGGAGGAGG 600  
 QY 601 TGCTGACAGAGCACCCTGTGAGCTGAGGCTGTCAGAGGACCTGTGAGCTGAGCC 660  
 DB 601 TGCTGACAGAGCACCCTGTGAGCTGAGGCTGTCAGAGGACCTGTGAGCTGAGCC 660  
 QY 661 TGCTGACAGAGCACCCTGTGAGCTGAGGCTGTCAGAGGACCTGTGAGCTGAGCC 720  
 DB 661 TGCTGACAGAGCACCCTGTGAGCTGAGGCTGTCAGAGGACCTGTGAGCTGAGCC 720  
 QY 721 TGCTACAGCCAGGCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 DB 721 TGCTACAGCCAGGCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 QY 781 TCAAGTAATTTCTTAGAAGATACATGAGGTTATTTCTTACTGATGAGAAATAGT 840  
 DB 781 TCAAGTAATTTCTTAGAAGATACATGAGGTTATTTCTTACTGATGAGAAATAGT 840  
 QY 841 GAAATCTAAGAAATTTTCTTTTCCAAAGCTATTTATTTCTGTGTATTCAGAGT 900  
 DB 841 GAAATCTAAGAAATTTTCTTTTCCAAAGCTATTTATTTCTGTGTATTCAGAGT 900  
 QY 901 ATTATCTCACTACATGATTTGTTGTGTAGGTTTCTTGAAGCTTAATTAATTTGA 960  
 DB 901 ATTATCTCACTACATGATTTGTTGTGTAGGTTTCTTGAAGCTTAATTAATTTGA 960  
 QY 961 AAAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1008  
 DB 961 AAAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1008

RESULT 2  
 ABS52633  
 ID ABS52633 standard; DNA; 1008 BP.

ABS52633;

15-NOV-2002 (first entry)

DNA encoding human secreted salivary protein zslg63.

Human; secreted salivary protein; zslg63; immunogen; zslg63-cytokine;  
 antibody-cytokine; in vivo killing; pathological microbe; bacteria;  
 fungal; viral; infection; salivary gland; anti-microbial; dental caries;  
 tooth decay; periodontal disease; thrush; gastrointestinal disease;  
 urinary tract infection; vaginal infection; skin infection; microflora;  
 epithelial wound; pathogenic colonisation; invasion; pro-inflammatory;  
 chronic tissue damage; vascular system; diabetes; anti-inflammatory;  
 incompetent immune system; AIDS; acquired immunodeficiency syndrome;  
 chemotherapy; radiation treatment; lung infection; cystic fibrosis;  
 digestion; chromosome 4; gene; db.

XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 128..787  
 FT /tag= a  
 FT /product= "Human secreted salivary protein zsig63"  
 FT misc\_feature 128..784  
 FT /tag= b  
 FT /note= "Specifically claimed in claim 3"  
 FT misc\_feature 173..784  
 FT /tag= c  
 FT /note= "Specifically claimed in claim 3"  
 FT US2002081701-A1.  
 XX  
 XX 27-JUN-2002.  
 XX  
 XX 03-AUG-2001; 2001US-0922480.  
 XX  
 XX 17-MAR-1999; 99US-124820P.  
 XX 17-MAR-2000; 2000US-0527345.  
 XX  
 XX (ADLER/) ADLER D A.  
 XX (SHEP/) SHEPPARD P O.  
 XX Adler DA, Sheppard PO;  
 XX WPI; 2002-635468/68.  
 XX P-PDB; ABG31608.  
 XX  
 XX Novel secreted salivary protein, zsig63 and polynucleotide encoding it  
 PT useful for treating microbial infections, inflammatory conditions, -  
 PT dental caries and lung infections associated with cystic fibrosis  
 XX  
 XX Claim 3; Page 27-28; 33pp; English.  
 XX  
 XX The present invention relates to a new secreted salivary protein, zsig63.  
 CC The invention is useful for detecting in a test sample, the presence of  
 CC an antagonist or agonist of zsig63 protein activity. The invention is  
 CC also useful as an immunogen for producing an antibody to zsig63  
 CC polypeptide. zsig63-cytokine fusion proteins or antibody-cytokine fusion  
 CC protein are useful for enhancing in vivo killing of target tissues.  
 CC Pharmaceutical composition comprising purified zsig63 polypeptide are  
 CC useful in the treatment of conditions associated with pathological  
 CC microbes, including bacterial, fungal and viral infections. High  
 CC expression of zsig63 in salivary gland suggests that anti-microbial  
 CC polypeptides are useful for treatment of dental caries (tooth decay),  
 CC periodontal disease, thrush and gastrointestinal disease. Other  
 CC applications can be used in urinary tract infections, vaginal infections,  
 CC prevention of infection in skin and other epithelial wounds. The  
 CC polypeptides can be used to establish normal microflora and protect  
 CC against pathogenic colonization and invasion. The invention is useful  
 CC when pro-inflammatory activity is desired. Applications for  
 CC damage, particularly in areas having a limited or damaged vascular system  
 CC e.g., damage in extremities associated with diabetes. Antagonists to  
 CC zsig63 polypeptides may be useful as anti-inflammatory agents. The  
 CC invention is useful for the treatment of patients having incompetent  
 CC immune system, such as AIDS (acquired immunodeficiency syndrome) patients  
 CC or individuals that have undergone chemotherapy, radiation treatment. The  
 CC invention is also useful for the treatment of lung infections associated  
 CC with cystic fibrosis and its agonists or antagonists are useful for  
 CC aiding digestion. The present nucleic acid sequence represents the human  
 CC zsig63 gene located on chromosome 4. This sequence encodes the human  
 CC secreted salivary protein zsig63 of the invention.  
 XX  
 XX Sequence 1008 BP; 257 A; 251 C; 193 G; 307 T; 0 other;

Query Match 100.0%; Score 1008; DB 24; Length 1008;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-207;  
 Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACAGACTTAAAAAGCCATGATTCCTTCCTCTCTAAAGAGAAAAATATTAATT 60  
 DB 1 AGACAGACTTAAAAAGCCATGATTCCTTCCTCTCTAAAGAGAAAAATATTAATT 60  
 QY 61 TAAAAATACATTCGCTATTTTCTTAAAAACAATAATTATAGCTTAATTCATAGGCTC 120  
 DB 61 TAAAAATACATTCGCTATTTTCTTAAAAACAATAATTATAGCTTAATTCATAGGCTC 120  
 QY 121 AATCAAAATGAAGCTTCCTTCCTTCGCGCTGACATTTGATGTGCTTTTGCAGAGAGAG 180  
 DB 121 AATCAAAATGAAGCTTCCTTCCTTCGCGCTGACATTTGATGTGCTTTTGCAGAGAGAG 180  
 QY 181 ACGGTTCCCTTCATTTGGTGAAGATGACATGACATGCGATCCCATTCATCTCTCT 240  
 DB 181 ACGGTTCCCTTCATTTGGTGAAGATGACATGACATGCGATCCCATTCATCTCTCT 240  
 QY 241 GAATATTCCTTATGCGATACGGAATTTACACCTCCTTTTATATGCGCCAGTGAATAC 300  
 DB 241 GAATATTCCTTATGCGATACGGAATTTACACCTCCTTTTATATGCGCCAGTGAATAC 300  
 QY 301 AGTCCCACTTACCTGGAATACCTTACACTGACACAGGGTTACCTTCGATCCCGAAT 360  
 DB 301 AGTCCCACTTACCTGGAATACCTTACACTGACACAGGGTTACCTTCGATCCCGAAT 360  
 QY 361 TCTAATCTCTCGATTCCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 420  
 DB 361 TCTAATCTCTCGATTCCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 420  
 QY 421 GTTGAATGTTCT 480  
 DB 421 GTTGAATGTTCT 480  
 QY 481 TTCAGAGCTGAGACACCGCTGCGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCT 540  
 DB 481 TTCAGAGCTGAGACACCGCTGCGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCT 540  
 QY 541 TCTTACAGCCACCTGTAAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 600  
 DB 541 TCTTACAGCCACCTGTAAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 600  
 QY 601 TGTGAGAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 660  
 DB 601 TGTGAGAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 660  
 QY 661 TGTGAGAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 720  
 DB 661 TGTGAGAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 720  
 QY 721 TGTACAGCCAAAGCTGCTGCGCCAGAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
 DB 721 TGTACAGCCAAAGCTGCTGCGCCAGAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
 QY 781 TCAAGTAATTTCTTGAAGAGTACCAAGGTTCAATTTCTATATGATGAGAAATAGT 840  
 DB 781 TCAAGTAATTTCTTGAAGAGTACCAAGGTTCAATTTCTATATGATGAGAAATAGT 840  
 QY 841 GAAATCTACAAAAGTTTCTTCTTCTTCCAAAAGACTTATTCATTCGTGTATTCAGAGT 900  
 DB 841 GAAATCTACAAAAGTTTCTTCTTCTTCCAAAAGACTTATTCATTCGTGTATTCAGAGT 900  
 QY 901 ATTCAATCTACATCATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960  
 DB 901 ATTCAATCTACATCATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960  
 QY 961 AAAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1008  
 DB 961 AAAAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1008

RESULT 3  
 AAS20591  
 ID AAS20591 standard; cDNA; 1008 BP.  
 XX

AC	AAS20591;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	Human zsi63 cDNA.
XX	
KW	Human; zsi63; chromosome 4q12-q413; salivary protein; antimicrobial; ss;
KM	microbial infection; tooth decay; periodontal disease; thrush; emphysema;
KV	gastrointestinal disease; urinary tract infection; vaginal infection;
KW	skin infection; epithelial wound; chronic tissue damage; cystic fibrosis;
KM	acquired immunodeficiency syndrome/AIDS; lung infection; sarcoidosis;
XX	chronic bronchitis; gene therapy; protein therapy; gene.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	128..787
FT	/tag= a
FT	/product= "Human zsi63"
PX	
PN	US6331413-B1.
PD	18-DEC-2001.
XX	
PF	17-MAR-2000; 2000US-0527345.
PR	17-MAR-1999; 99US-124820P.
PA	(ZYMO ) ZYMOGENETICS INC.
PI	Adler DA, Sheppard FO;
DR	WP1; 2002-096707/13.
DR	P-PSDB; AAU74536.
PT	Polynucleotides encoding salivary proteins useful as anti-microbial agents -
XX	
XX	Claim 2; Column 47-50; 29pp; English.
CC	The invention relates to a polynucleotide derived from the 4q12-q413
CC	region of human chromosome 4 and encoding a zsi63 polypeptide, a
CC	secreted salivary protein with anti-microbial activity. Due to their
CC	microbial activity, the sequences can be used in the study of microbial
CC	infections, e.g. For recombinant production of anti-microbial proteins.
CC	The sequences can be used in the treatment of tooth decay, periodontal
CC	disease, thrush, gastrointestinal disease, urinary tract infections,
CC	vaginal infections, skin infections, epithelial wounds, chronic tissue
CC	damage, acquired immunodeficiency syndrome (AIDS), cystic fibrosis, lung
CC	infections, sarcoidosis, emphysema and chronic bronchitis. This sequence
CC	represents cDNA encoding human zsi63.
SQ	
	Sequence 1008 BP; 257 A; 251 C; 193 G; 307 T; 0 other;
	Query Match 100.0%; Score 1008; DB 24; Length 1008;
	Best Local Similarity 100.0%; Pred. No. 3.5e-207;
	Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 AGACGACTAAAAAGCAGTATTCTTTGGTTCTCTTAAGAAGAAAATATAATT 60
Db	1 AGACGACTAAAAAGCAGTATTCTTTGGTTCTCTTAAGAAGAAAATATAATT 60
OY	TAAATAATCATTCGTATTTTCTTAAACAATAATTAAAGGTAAATTCATAGGGCT 120
Db	61 TAAATAATCATTCGTATTTTCTTAAACAATAATTAAAGGTAAATTCATAGGGCT 120
OY	121 AATCAAAATGAAGCTTCTTCCTTTGGGCCCTGCATTGATGTGTTCCTTTGCAAGGAAG 180
Db	121 AATCAAAATGAAGCTTCTTCCTTTGGGCCCTGCATTGATGTGTTCCTTTGCAAGGAAG 180
OY	181 ACGGTCCCCTTCATTGTGGAGAGATGACATGATGCTCACCCACTTCATCCATCTCT 240
Db	181 ACGGTCCCCTTCATTGTGGAGAGATGACATGATGCTCACCCACTTCATCCATCTCT 240

[illegible]

XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 128..787  
 FT /tag= a  
 FT /product= "zsig63"  
 FT /note= "this CDS (minus the stop codon) is claimed in  
 FT claim 3"  
 FT 128..172  
 FT /tag= b  
 FT 173..784  
 FT /tag= c  
 FT /label= Mature zsig63  
 FT /note= "this region is claimed in claim 3"  
 XX  
 XX US2002173027-A1.  
 XX  
 XX 21-NOV-2002.  
 XX  
 XX 03-AUG-2001; 2001US-0922469.  
 XX  
 XX 17-MAR-1999; 99US-124820P.  
 XX 17-MAR-2000; 2000US-0527345.  
 XX  
 XX (ADLER/) ADLER D A.  
 XX (SHEP/) SHEPPARD P O.  
 XX  
 XX Adler DA, Sheppard PO,  
 XX  
 XX MPI; 2003-328428/31.  
 XX P-Patent; AB08515.  
 XX  
 XX Novel isolated zsig63 polypeptide, member of the adhesin family, useful  
 XX for treating dental caries, periodontal disease, thrush,  
 XX gastrointestinal disease, urinary tract infections, vaginal infections,  
 XX skin infections -  
 XX  
 XX Claim 3; Page 26-27; 32pp; English.  
 XX  
 XX The invention relates to an isolated zsig63 polypeptide comprising at  
 XX least 90% identity to an amino acid sequence which comprises domain 1 of  
 XX zsig63, domain 2, domain 3, mature zsig63 and full length zsig63. Also  
 XX included are the polynucleotide encoding zsig63, a zsig63 expression  
 XX vector, a cultured cell comprising the vector and expressing the protein,  
 XX a DNA encoding a fusion protein (comprising amino acids 1-15, 16-37,  
 XX 38-126, 127-219 or 16-219 of zsig63 and an additional protein), using a  
 XX zsig63 reporter gene construct to identify zsig63 agonists, and  
 XX producing an anti-zsig63 antibody using zsig63 immunogenic peptides,  
 XX zsig63 is useful for detecting in a test sample, the presence of  
 XX antagonist of zsig63 protein activity. Zsig63 has antimicrobial  
 XX activity and since exhibits high expression in salivary gland, can be  
 XX used for treating dental caries, periodontal disease, thrush, and  
 XX gastrointestinal disease, urinary tract infections, vaginal infections,  
 XX skin infections and other epithelial wounds. The polypeptides can be  
 XX used to establish normal microflora and protect against pathogenic  
 XX colonization and invasion. Zsig63 can also be used for providing  
 XX pro-inflammatory activity for treating chronic, tissue damage  
 XX particularly in areas having limited or damaged vascular system, e.g.  
 XX in diabetes, and for treating immunocompromised AIDS patients or in  
 XX individuals that have undergone chemotherapy, radiation treatment, for  
 XX treating lung infections e.g. in cystic fibrosis. Detection of zsig63  
 XX polypeptide at relatively high levels in the trachea may indicate that  
 XX such polypeptides may serve as a marker of lung dysfunction. Zsig63 is  
 XX also useful in diagnosing conditions associated with salivary gland or  
 XX lung dysfunction including salivary gland carcinoma, Pneumocystis carinii  
 XX infection, emphysema, chronic bronchitis, prostate dysfunctions such  
 XX as prostate adenocarcinoma, aiding digestion, and as components of  
 XX defined cell culture media and may be used to replace serum that is  
 XX commonly used in culture. The DNA is useful in gene therapy applications  
 XX to increase or inhibit zsig63 activity, and for detecting abnormalities  
 XX on human chromosome 4 (e.g. 4q12-4q13, associated with dentinogenesis  
 XX imperfecta, and dentin dysplasia type II). Zsig63 is an adhesin family

CC member. The present sequence encodes human zsig63.  
 XX  
 XX Sequence 1008 BP; 257 A; 251 C; 193 G; 307 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 1008; DB 25; Length 1008;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-207;  
 Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 AGACGACGCTAAAGGCGGATGATCTTCTGCTCTTAAGAAAGAAATTAAT 60  
 1 AGACGACGCTAAAGGCGGATGATCTTCTGCTCTTAAGAAAGAAATTAAT 60  
 61 TAAATATACATGCGGATTTCTTAAACCAATTAATTATAGTTATATATAGGTC 120  
 61 TAAATATACATGCGGATTTCTTAAACCAATTAATTATAGTTATATATAGGTC 120  
 121 AATCAAAATGAAGCTTCTCTTGGGCGTGATGATGTTGCTTTGCAAGAGAG 180  
 121 AATCAAAATGAAGCTTCTCTTGGGCGTGATGATGTTGCTTTGCAAGAGAG 180  
 181 ACGGTTCCCTTCATTTGGGAGATGACAAATGATGTCACCACTTCATCTCT 240  
 181 ACGGTTCCCTTCATTTGGGAGATGACAAATGATGTCACCACTTCATCTCT 240  
 241 GAATATTCCTTATGCGAATATTAACCACTCTCTTATATATGCGGAGATATC 300  
 241 GAATATTCCTTATGCGAATATTAACCACTCTCTTATATATGCGGAGATATC 300  
 301 AGTCCCGAGTTACCTTGGGAAATCTTACATGACACAGGGTATCTTGATCTGAT 360  
 301 AGTCCCGAGTTACCTTGGGAAATCTTACATGACACAGGGTATCTTGATCTGAT 360  
 361 TCTAATCTCTCTGATATCCCTATGATGATGATGATGATGATGATGATGAT 420  
 361 TCTAATCTCTCTGATATCCCTATGATGATGATGATGATGATGATGATGAT 420  
 421 GTTGAATGTTCT 480  
 421 GTTGAATGTTCT 480  
 481 TTCAGCAGCTGACGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 481 TTCAGCAGCTGACGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 541 TCTTACAGCAGCAGCTGATGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 600  
 541 TCTTACAGCAGCAGCTGATGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 600  
 601 TCTGCAAGAGGACCTGTTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 660  
 601 TCTGCAAGAGGACCTGTTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 660  
 661 TCTGCAAGAGGACCTGTTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 720  
 661 TCTGCAAGAGGACCTGTTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 720  
 721 TCTGCAAGAGGACCTGTTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 780  
 721 TCTGCAAGAGGACCTGTTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 780  
 781 TCAAGTGAATTTCTTGAAGAGGATCAAGAGGATTTCTTGAAGAGGATTT 840  
 781 TCAAGTGAATTTCTTGAAGAGGATCAAGAGGATTTCTTGAAGAGGATTT 840  
 841 GAAATCTCAAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
 841 GAAATCTCAAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
 901 ATTATCTCAACATGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960  
 901 ATTATCTCAACATGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960  
 961 AAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1008





```

FT      /tag= a
XT      /product= "secreted protein clone cp16 1"
XX
XX      MO9901466-A1.
XX
XX      14-JAN-1999.
XX
XX      01-JUL-1998; 98WO-US13813.
XX
XX      27-OCT-1997; 97US-0958304.
XX      02-JUL-1997; 97US-0887195.
XX      (GEMT) GENETICS INST INC.
XX
XX      Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Racie LA;
XX      Spaulding V, Treacy M;
XX      WPI; 1999-105994/09.
XX      P-PSDB; AAW30653.
XX
XX      New polynucleotides encoding secreted human proteins - are derived
XX      from human foetal brain, adult testes, adult brain, foetal kidney,
XX      adult salivary gland, or adult blood cDNA libraries, useful as, e.g.
XX      potential vaccines
XX
XX      Claim 23; Page 70-71; 107bp; English.
XX
XX      The present sequence encodes a human secreted protein from clone
XX      CP16 1, deposited as ATCC 98482. Human secreted protein clone
XX      polynucleotides and proteins are predicted to have biological
XX      activities which would make them suitable for treating, preventing or
XX      ameliorating medical conditions in humans and animals. Suggested
XX      activities include nutritional activity, cytokine and cell
XX      proliferation/differentiation activity, immune stimulating (e.g. as
XX      vaccines) or suppressing activity, haematopoiesis regulating activity,
XX      tissue growth activity, activin/inhibin activity, chemotactic/
XX      chemokinetic activity, haemostatic and thrombolytic activity, receptor/
XX      ligand activity, anti-inflammatory activity, cadherin/tumour invasion
XX      suppressor activity, and tumour inhibition activity. The polynucleotides
XX      are also stated to be useful for gene therapy.
XX
XX      Sequence 1325 BP; 407 A; 289 C; 241 G; 388 T; 0 other;
XX
Query Match      97.4%; Score 981.6; DB 20; Length 1325;
Best Local Similarity 99.5%; Pred. No.1.9e-201;
Matches 995; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
XX
QY      10 AAAAAAGCCATGATCTTCTGTTCTCTTAAAGAGAAAAATATTAATTAATAATAC 69
DB      9 AAAAAAGCCATGATCTTCTGTTCTCTTAAAGAGAAAAATATTAATTAATAATAC 68
QY      70 ATGCGGATTTTCTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 129
DB      69 ATGCGGATTTTCTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 128
QY      130 GAAGCTTCTCTTGGGCTGATTTGATGTTGTTTGAAGAGAGAGAGAGAGAGAGAGAG 189
DB      129 GAAGCTTCTCTTGGGCTGATTTGATGTTGTTTGAAGAGAGAGAGAGAGAGAGAGAG 188
QY      190 CTTCATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249
DB      189 CTTCATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248
QY      250 TTATGGCATTCGGAATTTACACCTCTTTATTAATTCGCGAGTGAATCACTCCCGAG 309
DB      249 TTATGGCATTCGGAATTTACACCTCTCTTTATTAATTCGCGAGTGAATCACTCCCGAG 308
QY      310 TTACCTTGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 369
DB      309 TTACCTTGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 368
QY      370 TCTTGATTTCCCTATGTCATCAATCCGTTGTTTCCCTTGAATCACTGATTAATGT 429
DB      TCTTGATTTCCCTATGTCATCAATCCGTTGTTTCCCTTGAATCACTGATTAATGT 429

```

```

DB      369 TCTTGATTTCCCTATGTCATCAATCCGTTGTTTCCCTTGAATCACTGATTAATGT 428
QY      430 TCTTGATTTCCCTATGTCATCAATCCGTTGTTTCCCTTGAATCACTGATTAATGT 429
DB      429 TCTTGATTTCCCTATGTCATCAATCCGTTGTTTCCCTTGAATCACTGATTAATGT 428
QY      490 TGCAGACCCGCTGCCCCACTAATTCAGCTGAGCTGCTGAGCTGCACTTTACAGC 549
DB      489 TGCAGACCCGCTGCCCCACTAATTCAGCTGAGCTGCTGAGCTGCACTTTACATC 548
QY      550 CACACTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
DB      549 CACACTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
QY      610 GGCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
DB      609 GGCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
QY      670 GGCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
DB      669 GGCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728
QY      730 CAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 789
DB      729 CAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 788
QY      790 TTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 849
DB      789 TTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 848
QY      850 AAAAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 909
DB      849 AAAAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 908
QY      910 ACTACATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 968
DB      909 ACTACATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 968
QY      969 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1008
DB      969 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1008

RESULT 7
AAK61352
ID      AAK61352 standard; cDNA; 959 BP.
XX
XX      AAK61352;
XX
XX      14-JUL-1999 (first entry)
XX
XX      DE      DNA encoding a human secreted protein.
XX
XX      Human secreted protein; cancer; tumour; neurodegenerative disorder;
XX      developmental abnormality; fetal deficiency; blood disorder; leukemia;
XX      immune system disease; autoimmune disease; hepatic disease;
XX      renal disease; lymphoma; inflammation; allergy; ischemic shock;
XX      Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
XX      obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
XX      lung disease; thymus disease; digestive disorder; endocrine disorder;
XX      infection; AIDS; ss.
XX
XX      Homo sapiens.
XX
XX      MO9922243-A1.
XX
XX      06-MAY-1999.
XX
XX      23-OCT-1998; 98WO-US22376.
XX
XX      24-OCT-1997; 97US-0063387.
XX      24-OCT-1997; 97US-0062784.
XX      24-OCT-1997; 97US-0063088.

```







PS Claim 4; Page 28-29; 33pp; English.

XX The present invention relates to a new secreted salivary protein, zsi963.  
 CC The invention is useful for detecting in a test sample, the presence of  
 CC an antagonist or agonist of zsi963 protein activity. The invention is  
 CC also useful as an immunogen for producing an antibody to zsi963  
 CC polypeptide. zsi963-cytokine fusion proteins or antibody-cytokine fusion  
 CC proteins are useful for enhancing in vivo killing of target tissues.  
 CC Pharmaceutical composition comprising purified zsi963 polypeptide are  
 CC useful in the treatment of conditions associated with pathological  
 CC microbes, including bacterial, fungal and viral infections. High  
 CC expression of zsi963 in salivary gland suggests that anti-microbial  
 CC polypeptides are useful for treatment of dental caries (tooth decay),  
 CC periodontal disease, thrush and gastrointestinal disease. Other  
 CC applications can be used in urinary tract infections, vaginal infections,  
 CC prevention of infection in skin and other epithelial wounds. The  
 CC polypeptides can be used to establish normal microflora and protect  
 CC against pathogenic colonization and invasion. The invention is useful  
 CC when pro-inflammatory activity is desired. Applications for  
 CC such pro-inflammatory activity include the treatment of chronic tissue  
 CC damage, particularly in areas having a limited or damaged vascular system  
 CC e.g., damage in extremities associated with diabetes. Antagonists to  
 CC zsi963 polypeptides may be useful as anti-inflammatory agents. The  
 CC invention is useful for the treatment of patients having incompetent  
 CC immune system, such as AIDS (acquired immunodeficiency syndrome) patients  
 CC or individuals that have undergone chemotherapy, radiation treatment. The  
 CC invention is also useful for the treatment of lung infections associated  
 CC with cystic fibrosis and its agonists or antagonists are useful for  
 CC aiding digestion. The present nucleic acid sequence encodes the human  
 CC secreted salivary protein zsi963 of the invention.

XX Sequence 657 BP; 72 A; 150 C; 114 G; 65 T; 256 other;

Query Match 45.6%; Score 460; DB 24; Length 657;  
 Best Local Similarity 61.0%; Pred. No. 2.2e-89;  
 Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

QY 128 ATGAGCTTCTGCTTGGGCTGATGTATGTGTTGCTTTGCAAGAAGAGGTTTC 187  
 DB 1 ATGAATYTYTNTYNTGGGCTGATGTATGTGTTGCTTTGCAAGAAGAGGTTTC 60  
 QY 188 CCTTATGTGTAGAGATGACAAATGACATGTCACCACTTCATCTCGAATAT 247  
 DB 61 CCNTTATGTGNGARAGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 248 CCTTATGTGTAGAGATGACAAATGACATGTCACCACTTCATCTCGAATAT 307  
 DB 121 CCTTATGTGTAGAGATGACAAATGACATGTCACCACTTCATCTCGAATAT 180  
 QY 308 AGTTACCTTGGGAAATCTTACCTGACACAGGTTTACCTTCCATCTCGAATAT 367  
 DB 181 WSNATACCGNGAATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 QY 368 TCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 427  
 DB 241 WSNCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 428 GTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487  
 DB 301 GTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 488 GCTGACAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547  
 DB 361 GCGGACGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 QY 548 GCGACACCTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607  
 DB 421 GCGACACCTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 608 GAGGACCTGTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667  
 DB 481 GAGGACCTGTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY 668 GAGGACCTGTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727  
 DB 541 GAGGACCTGTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 728 GCGACCTGTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784  
 DB 601 GCGACCTGTGTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657

RESULT 10

AAS20592  
 ID AAS20592 standard; cDNA; 657 BP.

AC AAS20592;

DT 23-APR-2002 (first entry)

DE Human zsi963 degenerate cDNA coding sequence.

KW Human; zsi963; chromosome 4q12-q13; salivary protein; antimicrobial; as;  
 KW microbial infection; tooth decay; periodontal disease; thrush; emphysema;  
 KW gastrointestinal disease; urinary tract infection; vaginal infection;  
 KW skin infection; epithelial wound; chronic tissue damage; cystic fibrosis;  
 KW acquired immunodeficiency syndrome; AIDS; lung infection; sarcoidosis;  
 KW chronic bronchitis; gene therapy; protein therapy; gene.

OS Homo sapiens.

PN US6311413-B1.

PD 18-DEC-2001.

PF 17-MAR-2000; 2000US-0527345.

PR 17-MAR-1999; 99US-124820P.

PA (ZYMO) ZYMOGENETICS INC.

PI Adler DA, Shepard PO;

PT WPI; 2002-096707/13.

PS Claim 3; Column 51-52; 29pp; English.

XX The invention relates to a polynucleotide derived from the 4q12-q13  
 XX region of human chromosome 4 and encoding a zsi963 polypeptide, a  
 XX secreted salivary protein with anti-microbial activity. Due to their  
 XX microbial activity, the sequences can be used in the study of microbial  
 XX infections, e.g. for recombinant production of anti-microbial proteins.  
 XX The sequences can be used in the treatment of tooth decay, periodontal  
 XX disease, thrush, gastrointestinal disease, urinary tract infections,  
 XX vaginal infections, skin infections, epithelial wounds, chronic tissue  
 XX damage, acquired immunodeficiency syndrome (AIDS), cystic fibrosis, lung  
 XX infections, sarcoidosis, emphysema and chronic bronchitis. This sequence  
 XX represents a degenerate cDNA encoding human zsi963.

XX Sequence 657 BP; 72 A; 150 C; 114 G; 65 T; 256 other;

Query Match 45.6%; Score 460; DB 24; Length 657;  
 Best Local Similarity 61.0%; Pred. No. 2.2e-89;  
 Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

QY 128 ATGAGCTTCTGCTTGGGCTGATGTATGTGTTGCTTTGCAAGAAGAGGTTTC 187  
 DB 1 ATGAATYTYTNTYNTGGGCTGATGTATGTGTTGCTTTGCAAGAAGAGGTTTC 60  
 QY 188 CCTTATGTGTAGAGATGACAAATGACATGTCACCACTTCATCTCGAATAT 247  
 DB 61 CCNTTATGTGNGARAGATGATGATGATGATGATGATGATGATGATGATGAT 120

```

QY 248 CCTATAGCATAGGAATTACCACTCTCTTATATGCGCCAGTGAATACAGTCCC 307.
DB 121 CCGTAAGGAAATGMAAAYTTCNCNCNCTTATATATGACGCTGTAATACGTCN 180
QY 308 AGTTACCCCTGGGAATCTTACACTGACAGAGGTTACCTTCGATCCCTGATTTCACT 367
DB 181 WSNATCCNGAANAATACNTAATACGATGACGATGATGATGATGATGATGATGAT 240
QY 368 TCTCCTGATTCCTCTATGATCATCATCCGATGATTCCTCTAGCTACTAGTGAAT 427
DB 241 WSNCCNGATTCCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 428 GTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487
DB 301 GTTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 488 GGTGACGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
DB 361 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 548 GCGACACCTGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 607
DB 421 GCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480
QY 608 GAGGACCTGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 667
DB 481 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 668 GAGGACCTGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 727
DB 541 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 728 GCGACGCTGATGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 784
DB 601 GCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 657

```

RESULT 11  
ABX93595  
ID ABX93595 standard; DNA; 657 BP.

AC ABX93595;  
XX  
XX  
DT 28-MAY-2003 (first entry)  
XX  
XX  
DE Human zsig63 degenerate sequence.  
XX  
XX  
KW de; zsig63; adhesin; salivary gland; dental carrier;  
KW periodontal disease; thrush; gastrointestinal disease; epithelial wound;  
KW urinary tract infection; vaginal infection; skin infection;  
KW pro-inflammatory; chronic tissue damage; vascular system; diabetes; AIDS;  
KW lung infection; cystic fibrosis; lung dysfunction; digestive;  
KW salivary gland carcinoma; Pneumocystis carinii infection; emphysema;  
KW chronic bronchitis; prostate dysfunction; prostate adenocarcinoma;  
KW cell culture media; gene therapy; human chromosome 4q12-4q13;  
KW dentinogenesis imperfecta; dentin dysplasia type II.  
XX  
OS Synthetic.  
XX  
PN US2002173027-A1.  
XX  
PD 21-NOV-2002.  
XX  
PF 03-AUG-2001; 2001US-0922469.  
XX  
PR 17-MAR-1999; 99US-124820P.  
XX 17-MAR-2000; 2000US-0527345.  
XX  
PA (ADLER/) ADLER D A.  
XX (SHEP/) SHEPPARD P O.  
XX  
PI Adler DA, Sheppard PO;

```

DR WP1; 2003-328428/31.
XX
XX Novel isolated zsig63 polypeptide, member of the adhesin family, useful
PT for treating dental carrier, periodontal disease, thrush,
PT gastrointestinal disease, urinary tract infections, vaginal infections,
PT skin infections
XX
XX Claim 4; Page 28; 32pp; English.
PS
XX The invention relates to an isolated zsig63 polypeptide comprising at
CC least 90% identity to an amino acid sequence which comprises domain 1 of
CC zsig63, domain 2, domain 3, mature zsig63 and full length zsig63. Also
CC included are the polynucleotide encoding zsig63, a zsig63 expression
CC vector, a cultured cell comprising the vector and expressing the protein,
CC a DNA encoding a fusion protein (comprising amino acids 1-15, 16-37,
CC 38-126, 127-219 or 16-219 of zsig63 and an additional protein), using a
CC zsig63 reporter gene construct to identify zsig63 agonists, and
CC producing an anti-zsig63 antibody using zsig63 immunogenic peptides,
CC antagonist of zsig63 protein activity. Zsig63 has antimicrobial
CC activity and since exhibits high expression in salivary gland, can be
CC used for treating dental carrier, periodontal disease, thrush, and
CC gastrointestinal disease, urinary tract infections, vaginal infections,
CC skin infections and other epithelial wounds. The polypeptides can be
CC used to establish normal microflora and protect against pathogenic
CC colonization and invasion. Zsig63 can also be used for providing
CC pro-inflammatory activity for treating chronic, tissue damage
CC particularly in areas having limited or damaged vascular system, e.g.
CC in diabetes, and for treating immunocompromised AIDS patients or in
CC individuals that have undergone chemotherapy, radiation treatment, for
CC treating lung infections e.g. in cystic fibrosis. Detection of zsig63
CC polypeptide at relatively high levels in the trachea may indicate that
CC such polypeptides may serve as a marker of lung dysfunction. Zsig63 is
CC also useful in diagnosing conditions associated with salivary gland or
CC lung dysfunction including salivary gland carcinoma, Pneumocystis carinii
CC infection, emphysema, chronic bronchitis, prostate dysfunction such
CC as prostate adenocarcinoma, aiding digestion, and as components of
CC defined cell culture media and may be used to replace serum that is
CC commonly used in culture. The DNA is useful in gene therapy applications
CC to increase or inhibit zsig63 activity, and for detecting abnormalities
CC on human chromosome 4 (e.g. 4q12-4q13), associated with dentinogenesis
CC imperfecta, and dentin dysplasia type II). Zsig63 is an adhesin family
CC member. The present sequence represents all possible sequences that
CC may encode zsig63.
XX
SQ Sequence 657 BP; 72 A; 150 C; 114 G; 65 T; 256 other;

```

Query Match 45.6%; Score 460; DB 25; Length 657;  
Best Local Similarity 61.0%; Pred. No. 2,2e-89;  
Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

```

QY 128 ATGAAGCTTCTCTCTGAGCGCGCATGATGATGATGATGATGATGATGATGATGAT 187
DB 1 ATGAAGATTTTATTTATGAGCGCGCATGATGATGATGATGATGATGATGATGATGAT 60
QY 188 CCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 247
DB 61 CCGTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 248 CCTATAGCATAGGAATTACCACTCTCTTATATGCGCCAGTGAATACAGTCCC 307
DB 121 CCGTAAGGAAATGMAAAYTTCNCNCNCTTATATATGACGCTGTAATACGTCN 180
QY 308 AGTTACCCCTGGGAATCTTACACTGACAGAGGTTACCTTCGATCCCTGATTTCACT 367
DB 181 WSNATCCNGAANAATACNTAATACGATGACGATGATGATGATGATGATGATGAT 240
QY 368 TCTCCTGATTCCTCTATGATCATCATCCGATGATTCCTCTAGCTACTAGTGAAT 427
DB 241 WSNCCNGATTCCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 428 GTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487

```

Qy	Dp
30. GTTNCNCNCTYTNCCNCCMNGAGNTTYCCNTTYYGTNCCNCCMMSMGTITTYTWMSSGNC	3.60
488 GCTGCAAGCAACCGGCTTCCGCCACTATATGTGAGGTGAGGCTGCTGCAAGCTCACTCTTACA	54.7
361 GCGNCGNCNCNCCNCCNCCNATHGCGNGCGNCGNCGNCGNCGNCCNYTNACN	4.20
548 GCCACACTGTAGCACTGAGCTGTGTGAGGGGCCCCCTGTTGCAAGCTGAGCTCTGCA	60.7
421 GCNACNCCNGTNGCNGCGNCGNCCGNGCGNGGAGNCCNGTNGCGNCGNCGNCGN	48.0
608 GAGGCACCTGTTGGACCTGAGCTGCTGTGCAAGGCACTGTGTCAGCTGAGCTCTGTGCA	66.7
481 GARGCNCNCTNGTNGANCGNARGCCNCGNCGNCGNCCNGTNGCNGCGNCGNCGNCGN	54.0
668 GAGGCACTGTTGAGTGTGAGGCAAGCTGTGAGAGAACTTCAACAGCTGAGCTCTGTACA	72.7
541 GAGCNCNCTNGTNGNCTNGARGCCNCGNCGNCGNCGNARGCCNMSNCCGCGNCGNCGN	60.0
728 GCCAAGCTGTGCCCCGAACTCACTCTTCTCTCTTTGAAACAGGCAATACAG	78.4
601 GCNARGCNGCNCNCCNARGCCNACAYCCNMSNCCNMSNYTNGARGCAAGCAAAACAR	65.7

## RESULT 12

ID AAV89597 standard; cDNA; 470 BP.  
 AC AAV89597;  
 XX  
 DT 15-FEB-1999 (first entry)  
 XX  
 DE EST clone CP251.  
 XX  
 KM Human; secreted protein; expressed sequence tag; EST; haematopoiesis;  
 KM tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;  
 KM receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
 KM gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN M09845436-A2.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98MO-US06955.  
 XX  
 PR 10-APR-1997; 97US-0838821.  
 XX  
 XX (GENY ) GENETICS INST INC.  
 PA  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
 PI Racine LA, Spaulding V, Treacy M;  
 XX  
 DR WPI; 1999-070077/06.  
 XX  
 PT New polynucleotides encoding human secreted proteins - derived from  
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PT ovary, pituitary, retina and colon cDNA libraries.  
 XX  
 XX  
 PS Claim 1; Page 269; 618pp; English.  
 XX  
 XX The present sequence represents a human expressed sequence tag (EST).  
 CC The polynucleotide, which is a secreted EST, and the encoded protein  
 CC are predicted to have useful biological activities which would make  
 CC them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is  
 CC given. Suggested activities include nutritional activity, immune  
 CC stimulating or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic  
 CC activity, receptor/ligand activity, anti-inflammatory activity,  
 CC cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 470 BP; 116 A; 115 C; 83 G; 156 T; 0 other;  
SQ

Query Match	44.5%	Score 449;	DB 20;	Length 470;
Best Local Similarity	98.9%	Pred. No. 4.5e-87;		
Matches 452; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

QY	11	AAAAAGCCATGATTCCTTCGTTTCTCTCTAAAGAGAAATAATATTTAAATAATACA	70
Db	11	AAGAAGCCATGATTCCTTCGTTTCTCTCTAAAGAGAAATAATATTTAAATAATACA	70
QY	71	TTGCGATTTTCTAAAAACATTAATTNAGTGTATATTCATAGGCTCATCAAAATG	130
Db	71	TTGCGATTTTCTAAAAACATTAATTNAGTGTATATTCATAGGCTCATCAAAATG	130
QY	131	AAGCTCTCTCTTGCGCCGCACTGTATAGTGTGCTTTGCAAGAGAAAGACGGTTCCCC	190
Db	131	AAGCTCTCTCTTGCGCCGCACTGTATAGTGTGCTTTGCAAGAGAAAGACGGTTCCCC	190
QY	191	TTGATTTGAGAGATGACATATGACATGATGTCACCCACTTCATTCATCTGTAATATTCCT	250
Db	191	TTGATTTGAGAGATGACATATGACATGATGTCACCCACTTCATTCATCTGTAATATTCCT	250
QY	251	TATGGCATACGGAAATTTACACCTCTCTTTATATATGCCCCAGAGATATACATGCCCAAGT	310
Db	251	TATGGCATACGGAAATTTACACCTCTCTTTATATATGCCCCAGAGATATACATGCCCAAGT	310
QY	311	TACCTCGGGAATACCTTACATGACACAGGGTTACCTTGATCCCTGGAATCTTCACTTCT	370
Db	311	TACCTCGGGAATACCTTACATGACACAGGGTTACCTTGATCCCTGGAATCTTCACTTCT	370
QY	371	CTGTGATTCGCCCTATATGCTATACATCCGTGGATTTCCCTTAGTACTCAAGTGAATGTT	430
Db	371	CTGTGATTCGCCCTATATGCTATACATCCGTGGATTTCCCTTAGTACTCAAGTGAATGTT	430
QY	431	CTCTCTCTCCCTCTTAGGGGTTTCCCGTTTGTCCTC	467
Db	431	CTCTCTCTCCCTCTTAGGGGTTTCCCGTTTGTCCTC	467

## RESULT 13

```

ID      AA089646 standard; cDNA; 450 BP.
XX
XX
AC      AAV89646;
XX
XX
DT      15-FEB-1999 (first entry)
XX
DE      EST clone CP116.
XX
XX      Human; secreted protein; expressed sequence tag; EST; hematopoiesis;
KW      tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
KW      receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
XX      gene therapy; ss.
XX
XX      Homo sapiens.
XX
XX      OS
XX      FN      WO9845436-A2.
XX
XX      PD      15-OCT-1998.
XX
XX      PF      10-APR-1998; 98WO-US06955.
XX
XX      PR      10-APR-1997; 97US-0838821.
XX
XX      (GEMV ) GENETICS INST INC.
XX
XX      PA
XX      PI      Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
XX      PI      Racie LA, Spaulding V, Treacy M;
XX      DR      WPI; 1999-070077/06.
XX
XX      New polynucleotides encoding human secreted proteins - derived from
PT

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:23:58 ; Search time 89 Seconds  
(without alignments)  
4999.036 Million cell updates/sec

Title: US-09-923-236-1

Perfect score: 1008

Sequence: 1 agacagactaaagacat.....tagataatcagaccatg 1008

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PTUS.COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008	100.0	1008	US-09-527-345-1	Sequence 1, Appli
2	460	45.6	657	US-09-527-345-3	Sequence 3, Appli
3	121.2	12.0	3489	US-08-728-333A-1	Sequence 1, Appli
4	121.2	12.0	3489	US-09-298-568-1	Sequence 1, Appli
5	121.2	12.0	3489	US-09-410-399-1	Sequence 1, Appli
6	121.2	12.0	32207	US-08-770-379-20	Sequence 20, Appli
7	121.2	12.0	32207	US-08-757-669A-20	Sequence 20, Appli
8	121.2	12.0	32207	US-09-230-371A-20	Sequence 20, Appli
9	100	9.9	1037	US-09-181-585-3	Sequence 3, Appli
10	100	9.9	1159	US-09-181-585-1	Sequence 1, Appli
11	100	9.9	1471	US-09-181-585-2	Sequence 2, Appli
12	93.6	9.3	397	US-09-253-691-3	Sequence 3, Appli
13	92	9.1	543	5273901-6	Patent No. 5273901
14	87.6	8.7	234	US-08-469-802B-3	Sequence 3, Appli
15	87.6	8.7	234	US-08-267-803B-3	Sequence 3, Appli
16	86.6	8.6	786	US-08-403-379A-2	Sequence 2, Appli
17	86.6	8.6	786	US-08-929-414-2	Sequence 2, Appli
18	85.2	8.5	1926	US-09-249-585A-2	Sequence 2, Appli
19	85.2	8.5	1926	US-09-410-399-3	Sequence 3, Appli
20	85.2	8.5	2580	US-09-050-863-2	Sequence 2, Appli
21	85.2	8.5	2580	US-09-359-081-2	Sequence 2, Appli
22	85.2	8.5	5452	US-09-130-114-1	Sequence 1, Appli
23	85.2	8.5	8705	US-09-647-344A-14	Sequence 14, Appli
24	85.2	8.5	9600	US-08-910-647-1	Sequence 1, Appli
25	85.2	8.5	9600	US-09-620-925-1	Sequence 1, Appli
26	85.2	8.5	10596	US-07-884-811-15	Sequence 15, Appli
27	85.2	8.5	10596	US-07-885-971-15	Sequence 15, Appli

C 28	85.2	8.5	10596	1	US-08-087-783A-15	Sequence 15, Appli
C 29	85.2	8.5	10596	2	US-08-194-088B-15	Sequence 15, Appli
C 30	85.2	8.5	10596	5	PCT-US93-04648-15	Sequence 15, Appli
C 31	85.2	8.5	10596	5	PCT-US93-04648-15	Sequence 1, Appli
C 32	84	8.3	1037	4	US-09-181-585-3	Sequence 3, Appli
C 33	84	8.3	1159	4	US-09-181-585-1	Sequence 1, Appli
C 34	84	8.3	1471	4	US-09-181-585-2	Sequence 2, Appli
C 35	83.6	8.3	477	4	US-09-135-994-1	Sequence 1, Appli
C 36	83.6	8.3	477	4	US-09-684-843A-1	Sequence 1, Appli
C 37	83.2	8.3	397	3	US-09-253-691-3	Sequence 1, Appli
C 38	83	8.2	2830	2	US-09-010-928B-1	Sequence 1, Appli
C 39	80	7.9	533	6	5482709-5	Patent No. 5482709
C 40	78.2	7.8	16442	3	US-08-781-891-208	Sequence 208, App
C 41	78.2	7.8	16442	4	US-09-618-166-208	Sequence 7, Appli
C 42	77.8	7.7	203	3	US-09-043-303-7	Sequence 2, Appli
C 43	77.2	7.7	2824	2	US-09-010-928B-3	Sequence 3, Appli
C 44	75.8	7.5	195	1	US-08-469-802B-2	Sequence 2, Appli
C 45	75.8	7.5	195	2	US-08-267-803B-2	Sequence 2, Appli

## ALIGNMENTS

```
RESULT 1
US-09-527-345-1
; Sequence 1, Application US/09527345
; Patent No. 6331413
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: SECRETED SALIVARY ZS1G63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/527,345
; PRIOR APPLICATION NUMBER: 1999-03-17
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)...(784)
US-09-527-345-1

Query Match      100.0%; Score 1008; DB 4; Length 1008;
Best Local Similarity 100.0%; Pred. No. 2.7e-234;
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACAGCTAAAGGCGATGATCTTCTGTTCTCTTAAAGAGAAATTAATT 60
   1 AGACAGCTAAAGGCGATGATCTTCTGTTCTCTTAAAGAGAAATTAATT 60
Db 1 AGACAGCTAAAGGCGATGATCTTCTGTTCTCTTAAAGAGAAATTAATT 60
QY TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 120
   TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 120
Db TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 120
QY TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 180
   TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 180
Db TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 180
QY TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 240
   TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 240
Db TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 240
QY TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 300
   TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 300
Db TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 300
QY TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 360
   TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 360
Db TAAATAATGAGGATTTCTTAAACATTAATTAATGTTAATATTCATAGGTC 360
```

```

Db      301 AGTCCCAAGTACCTCGGAAATCTTACACTGACAGAGGGTACTCTGATCCCTGAT 360
Qy      361 TCTAATCTCCGGAATCCCGATNGTATGACATCCGGTGTTCCTTAGCTACTCA 420
Db      361 TCTAATCTCCGGAATCCCGATNGTATGACATCCGGTGTTCCTTAGCTACTCA 420
Qy      421 GTTGAATGTTCCCTCTCTCCCTCTAGGGGTTCCCGTTTGTCCCTCTCAAGTTT 480
Db      421 GTTGAATGTTCCCTCTCTCCCTCTAGGGGTTCCCGTTTGTCCCTCTCAAGTTT 480
Qy      481 TTGAGAGCTGAGACACCCGCTGCCACCTATTGACAGCTGAGCTGAGCTGAGCC 540
Db      481 TTGAGAGCTGAGACACCCGCTGCCACCTATTGACAGCTGAGCTGAGCTGAGCC 540
Qy      541 TCTTACAGGCAACCTGTAGAGCTGAGCTGAGCTGAGGGGCCCTGTGAGCTGAGCC 600
Db      541 TCTTACAGGCAACCTGTAGAGCTGAGCTGAGCTGAGGGGCCCTGTGAGCTGAGCC 600
Qy      601 TGTGAGAGGCACTGTGAGCTGAGCTGAGCTGAGAGGCACTGTGAGCTGAGCC 660
Db      601 TGTGAGAGGCACTGTGAGCTGAGCTGAGCTGAGAGGCACTGTGAGCTGAGCC 660
Qy      661 TGTGAGAGGCACTGTGAGCTGAGCTGAGCTGAGAGGCACTGTGAGCTGAGCC 720
Db      661 TGTGAGAGGCACTGTGAGCTGAGCTGAGCTGAGAGGCACTGTGAGCTGAGCC 720
Qy      721 TGTGAGAGGCACTGTGAGCTGAGCTGAGCTGAGAGGCACTGTGAGCTGAGCC 780
Db      721 TGTGAGAGGCACTGTGAGCTGAGCTGAGCTGAGAGGCACTGTGAGCTGAGCC 780
Qy      781 TGAAGTAAATCTTGAAGAGTACCATGGTTCATTTGATATGATGAGTAAAT 840
Db      781 TGAAGTAAATCTTGAAGAGTACCATGGTTCATTTGATATGATGAGTAAAT 840
Qy      841 GAAATCTACAAAGTTTCTTTCTTTCCAAAGACTATTTCACTGTGTGATGAGT 900
Db      841 GAAATCTACAAAGTTTCTTTCTTTCCAAAGACTATTTCACTGTGTGATGAGT 900
Qy      901 ATTCAATCTACATGATGATTTGTTGTGTGATTTTCTTGGACTTAATTATTTGA 960
Db      901 ATTCAATCTACATGATGATTTGTTGTGTGATTTTCTTGGACTTAATTATTTGA 960
Qy      961 AAAAATGATGATTAATTAATAATAATAGATTAATTAGAACCAATGG 1008
Db      961 AAAAATGATGATTAATTAATAATAATAGATTAATTAGAACCAATGG 1008

RESULT 2
US-09-527-345-3
; Sequence 3, Application US/09527345
; Patent No. 631413
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/527,345
; CURRENT FILING DATE: 1999-03-17
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence for zsig63
; NAME/KEY: misc. feature
; LOCATION: (1)...(657)
; OTHER INFORMATION: n = A,T,C or G
US-09-527-345-3

```

```

Query Match      45.6%; Score 460; DB 4; Length 657;
Best Local Similarity 61.0%; Pred. No. 5 6e-102;
Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

Qy      128 ATGAAGCTTCTCCCTTGGGCTGATATGATGTTGCTTTTGAAGAAGACGGTTC 187
Db      1  ATGAATATATATATGAGGCTGATATGATGTTGCTTTTGAAGAAGACGGTTC 60
Qy      188 CCTTCATTTGGAGAGATGACATGATGATGATGATGATGATGATGATGATGATGAT 247
Db      61  CCTTATATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy      248 CCTTATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 307
Db      121 CCTTATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Qy      308 AGTTACCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 367
Db      181 WNTATVCCGNAAYACNTATVACNTATVACNTATVACNTATVACNTATVACNTAT 240
Qy      368 TCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 427
Db      241 WNTATVCCGNAAYACNTATVACNTATVACNTATVACNTATVACNTATVACNTAT 300
Qy      428 GTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487
Db      301 GTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Qy      488 GCTGAGAGACCCGCTGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCTTACA 547
Db      361 GCTGAGAGACCCGCTGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCTTACA 420
Qy      548 GCTGAGAGACCCGCTGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCTTACA 607
Db      421 GCTGAGAGACCCGCTGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCTTACA 480
Qy      608 GAGGACCTGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 667
Db      481 GAGGACCTGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
Qy      668 GAGGACCTGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 727
Db      541 GAGGACCTGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 600
Qy      728 GCTGAGAGACCCGCTGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCTTACA 784
Db      601 GCTGAGAGACCCGCTGCCCACTTATGAGCTGAGCTGAGCTGAGCTGAGCTTACA 657

RESULT 3
US-08-728-323A-1/c
; Sequence 1, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Belman, Isidore S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/728,323A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSC/SKS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3489 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3489  
US-08-728-323A-1

Query Match 12.0%; Score 121.2; DB 2; Length 3489;  
Best Local Similarity 58.7%; Pred. No. 8.9e-20;  
Matches 210; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 413 GCTACTCAAGTGAATGTCCTCTCCCTCCCTCCTAGAGGGTTCCCGTTGCTCCTCTCA 472  
DB 2188 GCTGCTCATCCCTGCTGCTGCTGCTCATCCCTGCTGCTGCTGCTGCTGCTGCTCAT 2129

QY 473 AGGTTTTTTCAGCACTGAGACACCCGCTGCCACCTATTCAGCTGAGCTGCTGCA 532  
DB 2128 CCGTGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069

QY 533 GCTGACCTCTTACAGACCAACCTGTAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCA 592  
DB 2068 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009

QY 593 GCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 652  
DB 2008 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949

QY 653 GCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 712  
DB 1948 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889

QY 713 GCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 770  
DB 1888 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831

RESULT 4  
US-09-298-568-1/c  
Sequence 1, Application US/09298568  
Patent No. 6332792  
GENERAL INFORMATION:  
APPLICANT: Kieff, Elliott D.  
APPLICANT: Kieff, Elliott D.  
APPLICANT: Kieff, Elliott D.  
APPLICANT: Kieff, Elliott D.  
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
FILE REFERENCE: 16412-10001R  
CURRENT APPLICATION NUMBER: US/09/298,568  
CURRENT FILING DATE: 1999-04-21  
EARLIER APPLICATION NUMBER: US 60/109,422  
EARLIER FILING DATE: 1998-11-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3489

TYPE: DNA  
ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-298-568-1

Query Match 12.0%; Score 121.2; DB 4; Length 3489;  
Best Local Similarity 58.7%; Pred. No. 8.9e-20;  
Matches 210; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 413 GCTACTCAAGTGAATGTCCTCTCCCTCCCTCCTAGAGGGTTCCCGTTGCTCCTCTCA 472  
DB 2188 GCTGCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCAT 2129

QY 473 AGGTTTTTTCAGCACTGAGACACCCGCTGCCACCTATTCAGCTGAGCTGCTGCA 532  
DB 2128 CCGTGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069

QY 533 GCTGACCTCTTACAGACCAACCTGTAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCA 592  
DB 2068 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009

QY 593 GCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 652  
DB 2008 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949

RESULT 5  
US-09-410-399-1/c  
Sequence 1, Application US/09410399  
Patent No. 6482587  
GENERAL INFORMATION:  
APPLICANT: Robertson, Eric S.  
APPLICANT: Robertson, Eric S.  
APPLICANT: Robertson, Eric S.  
APPLICANT: Robertson, Eric S.  
TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA  
FILE REFERENCE: UM-03778  
CURRENT APPLICATION NUMBER: US/09/410,399  
CURRENT FILING DATE: 1999-10-01  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3489  
TYPE: DNA  
ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-410-399-1

Query Match 12.0%; Score 121.2; DB 4; Length 3489;  
Best Local Similarity 58.7%; Pred. No. 8.9e-20;  
Matches 210; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 413 GCTACTCAAGTGAATGTCCTCTCCCTCCCTCCTAGAGGGTTCCCGTTGCTCCTCTCA 472  
DB 2188 GCTGCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCAT 2129

QY 473 AGGTTTTTTCAGCACTGAGACACCCGCTGCCACCTATTCAGCTGAGCTGCTGCA 532  
DB 2128 CCGTGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069

QY 533 GCTGACCTCTTACAGACCAACCTGTAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCA 592  
DB 2068 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009

QY 593 GCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 652  
DB 2008 GCTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949

QY 653 GCTGAGCCTGCTGACAGAGCACTGTTGAGTGGAGCCAGCTGACAGAGAACTTACCA 712  
DB 1948 GCTCATCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCT 1889  
QY 713 GCTGAGCCTGCTGACAGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770  
DB 1888 GCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831

## RESULT 6

US-08-770-379-20  
Sequence 20, Application US/08770379  
Patent No. 5849564

GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770.379  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 12.0%; Score 121.2; DB 2; Length 32207;  
Best Local Similarity 58.7%; Pred. No. 2.6e-19;  
Matches 210; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 413 GCTACTCAGTTGAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 472  
DB 19809 GCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCTGCTCAT 19868  
QY 473 AGGTTTTTTTACAGAGCTGACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532  
DB 19869 CCGTGTCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19928  
QY 533 GCTGACCTCTTACAGCCACACCTGTAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGA 592  
DB 19929 GCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCT 19988  
QY 593 GCTGAGCCTGCTGACAGAGCACTGTTGAGTGGAGCCAGCCTGCTGCTGCTGCTGCTGCTGA 652  
DB 19989 GCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCT 20048

QY 653 GCTGAGCCTGCTGACAGAGCACTGTTGAGTGGAGCCAGCCTGACAGAGAACTTACCA 712  
DB 20049 GCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCT 20108  
QY 713 GCTGAGCCTGCTGACAGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770  
DB 20109 GCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCT 20166

## RESULT 7

US-08-757-669A-20  
Sequence 20, Application US/08757669A  
Patent No. 6183751

GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/757.669A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 45185-F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-757-669A-20

Query Match 12.0%; Score 121.2; DB 3; Length 32207;  
Best Local Similarity 58.7%; Pred. No. 2.6e-19;  
Matches 210; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 413 GCTACTCAGTTGAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 472  
DB 19809 GCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCTGCTCAT 19868  
QY 473 AGGTTTTTTTACAGAGCTGACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532  
DB 19869 CCGTGTCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19928  
QY 533 GCTGACCTCTTACAGCCACACCTGTAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGA 592  
DB 19929 GCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCT 19988  
QY 593 GCTGAGCCTGCTGACAGAGCACTGTTGAGTGGAGCCAGCCTGCTGCTGCTGCTGCTGCTGA 652

Db 19989 GCTCATCTGCTGCTGCTCATCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCT 20048  
Qy 653 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712  
Db 20049 GCTCATCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCT 20108  
Qy 713 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770  
Db 20109 GCTCATCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCTGCTCATCTGCTGCT 20166

## RESULT 8

US-09-230-371A-20

; Sequence 20, Application US/09230371A  
; Patent No. 6348586  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A  
; APPLICANT: Russco, James J  
; APPLICANT: Edelman, Isidore S  
; APPLICANT: Moore, Patrick S  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
; FILE REFERENCE: 45185-G-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,371A  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR FILING DATE: 1997-07-22  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 32207  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-230-371A-20

Query Match 12.0%; Score 121.2; DB 4; Length 32207;  
Best Local Similarity 58.7%; Pred. No. 2.6e-19;  
Matches 210; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy 413 GCTACTCAGTTGATGTTCA 472  
Db 19809 GCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCA 19868  
Qy 473 AGGTTTTTTCAGAGCTGAGACCGCGTGGCCCACTATTGACGTGAGCTGCTGCA 532  
Db 19869 CCGTCTGCTCTCTCTCATCTGCTGCTCTCTCTCATCTGCTGCTGCTGCTGCTGCT 19928  
Qy 533 GCTGCACTCTTTCAGAGCACTGTAAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCA 592  
Db 19929 GCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19988  
Qy 593 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652  
Db 19989 GCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20048  
Qy 653 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712  
Db 20049 GCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20108  
Qy 713 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770  
Db 20109 GCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20166

## RESULT 9

US-09-181-585-3

; Sequence 3, Application US/09181585  
; Patent No. 6524791  
; GENERAL INFORMATION:  
; APPLICANT: Rannum, Laura P.W.  
; APPLICANT: Koob, Michael  
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION

; FILE REFERENCE: 1100900101  
; CURRENT APPLICATION NUMBER: US/09/181,585  
; CURRENT FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1037  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: cDNA  
US-09-181-585-3

Query Match 9.9%; Score 100; DB 4; Length 1037;  
Best Local Similarity 56.8%; Pred. No. 6.6e-15;  
Matches 184; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 406 TCCCTTAGCTACTCAGTTGATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 465  
Db 591 TTCCTTAGCTAGACCTGGGCTCTCTCATGTTAGAAAACCTGGCTTACTACTACTACTAC 650  
Qy 466 TCCCTTAGCTTTTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525  
Db 651 TACT 710  
Qy 526 TCCCTGAGCTGACCTCTTTCAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585  
Db 711 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770  
Qy 586 TCTTGCAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645  
Db 771 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830  
Qy 646 TCTTGCAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705  
Db 831 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890  
Qy 706 TTCACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
Db 891 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914

## RESULT 10

US-09-181-585-1

; Sequence 1, Application US/09181585  
; Patent No. 6524791  
; GENERAL INFORMATION:  
; APPLICANT: Rannum, Laura P.W.  
; APPLICANT: Koob, Michael  
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION  
; FILE REFERENCE: 1100900101  
; CURRENT APPLICATION NUMBER: US/09/181,585  
; CURRENT FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1159  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-181-585-1

Query Match 9.9%; Score 100; DB 4; Length 1159;  
Best Local Similarity 56.8%; Pred. No. 7e-15;  
Matches 184; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 406 TCCCTTAGCTACTCAGTTGATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 465  
Db 403 TTCCTTAGCTAGACCTGGGCTCTCTCATGTTAGAAAACCTGGCTTACTACTACTACTAC 462  
Qy 466 TCCCTTAGCTTTTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525  
Db 463 TACT 522



QY 550 CACACCTGTGAGAGCTGCTGCTGAGGAGCCCTGTGTGAGCTGAGCTGCTGAGCA 609  
DB 234 TGTCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175  
QY 610 GGCACCTGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669  
DB 174 TGTCTGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 115  
QY 670 GGCACCTGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
DB 114 TGTCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 55  
QY 730 CAGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769  
DB 54 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 15

## RESULT 14

US-08-469-802B-3/c  
Sequence 3, Application US/08469802B  
Patent No. 5741645  
GENERAL INFORMATION:  
APPLICANT: Orr, Harry T.  
APPLICANT: Rannum, Laura P.W.  
APPLICANT: Chung, Ming-Yi  
APPLICANT: Zoghbi, Huda Y.  
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
Patent No. 5741645

TITLE OF INVENTION: Type 1 and Method for Diagnosis  
NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.  
STREET: 119 No. 5741645th Fourth Street, Suite 203

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,802B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mueeling, Ann M.

REGISTRATION NUMBER: 33,977

REFERENCE/DOCKET NUMBER: 110.00030101

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-305-1217

TELEFAX: 612-305-1225

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 234 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-469-802B-3

Query Match

Best Local Similarity 63.1%; Score 87.6; DB 1; Length 234;

Matches 135; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 482 TCAGCAGCTGACGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541  
DB 214 TAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155  
QY 542 CTTAAGCAGCAACCTGTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601  
DB 154 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 95

QY 602 GCTGAGAGGACCTGTTGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661  
DB 94 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 35  
QY 662 GCTGAGAGGACCTGTTGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695  
DB 34 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1

## RESULT 15

US-08-267-803B-3/c  
Sequence 3, Application US/08267803B  
Patent No. 5834183  
GENERAL INFORMATION:  
APPLICANT: Orr, Harry T.  
APPLICANT: Rannum, Laura P.W.  
APPLICANT: Chung, Ming-Yi  
APPLICANT: Zoghbi, Huda Y.  
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia  
Patent No. 5834183

TITLE OF INVENTION: Type 1 and Method for Diagnosis  
NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.  
STREET: P.O. Box 581415

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55458-1415

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/267,803B

FILING DATE: 28-JUN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McCormack, Myra H.

REGISTRATION NUMBER: 36,602

REFERENCE/DOCKET NUMBER: 110.00030120

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-305-1217

TELEFAX: 612-305-1228

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 234 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-267-803B-3

Query Match

Best Local Similarity 63.1%; Score 87.6; DB 2; Length 234;

Matches 135; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 482 TCAGCAGCTGACGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541  
DB 214 TAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155  
QY 542 CTTAAGCAGCAACCTGTGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601  
DB 154 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 95  
QY 602 GCTGAGAGGACCTGTTGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661  
DB 94 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 35  
QY 662 GCTGAGAGGACCTGTTGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695



Wed Jan 21 09:54:18 2004

us-09-923-236-1.rn1

Page 8

Db 34 GCTGCTGCTGCTGCTGCTGCTGCTGCTG 1

Search completed: January 16, 2004, 09:24:29  
Job time : 92 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 08:42:28 ; Search time 392 Seconds  
(without alignments)  
9063.677 Million cell updates/sec

Title: US-09-923-236-1  
Perfect score: 1008  
Sequence: 1 agacagacacacacacacat.....tagataattagacacatg 1008

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008	100.0	1008	9	US-09-922-480-1
2	1008	100.0	1008	9	US-09-923-236-1
3	1008	100.0	1008	10	US-09-922-469-1
4	981.6	97.4	1325	11	US-09-746-783-105
5	896.4	88.9	959	13	US-10-231-417-41
6	460	45.6	657	9	US-09-922-480-3
7	460	45.6	657	9	US-09-923-236-3
8	460	45.6	657	10	US-09-922-469-3
9	121.2	12.0	3489	13	US-10-294-804-1
10	100	9.9	1037	12	US-10-373-667-3
11	100	9.9	1159	12	US-10-373-667-1
12	100	9.9	1471	12	US-10-373-667-2
13	98.2	9.7	2108	10	US-09-962-832-225
14	92.4	9.2	1101	9	US-09-874-062-2
15	89.2	8.8	2183	12	US-10-104-047-1064

C 16	87.6	8.7	1101	9	US-09-874-062-2	Sequence 2, Appl1
C 17	85.2	8.5	1262	13	US-10-294-804-3	Sequence 3, Appl1
C 18	85.2	8.5	8705	15	US-10-291-230-14	Sequence 14, Appl1
C 19	85.2	8.5	8705	15	US-10-291-249-14	Sequence 14, Appl1
C 20	85.2	8.5	9600	13	US-10-278-751-1	Sequence 1, Appl1
C 21	85.2	8.5	10233	13	US-10-050-898-283	Sequence 283, App
C 22	85.2	8.5	10285	13	US-10-050-902-283	Sequence 283, App
C 23	84	8.3	1037	12	US-10-373-667-3	Sequence 3, Appl1
C 24	84	8.3	1159	12	US-10-373-667-1	Sequence 1, Appl1
C 25	84	8.3	1471	12	US-10-373-667-2	Sequence 2, Appl1
C 26	83.6	8.3	477	13	US-10-357-322-1	Sequence 1, Appl1
C 27	83	8.2	2183	12	US-10-104-047-1064	Sequence 1064, App
C 28	78.2	7.8	16442	12	US-10-374-077-208	Sequence 208, App
C 29	77.8	7.7	422	10	US-09-738-977-337	Sequence 337, App
C 30	77.8	7.7	422	10	US-09-854-133-337	Sequence 337, App
C 31	77.8	7.7	422	15	US-10-144-644A-337	Sequence 337, App
C 32	75	7.4	14859	12	US-10-051-874-55	Sequence 55, Appl
C 33	73.4	7.3	9369	13	US-10-200-562-190	Sequence 190, App
C 34	73.4	7.3	9369	13	US-10-237-551-190	Sequence 190, App
C 35	73.4	7.3	9369	13	US-10-237-551-247	Sequence 247, App
C 36	73.4	7.3	15476	13	US-09-827-688-8	Sequence 8, Appl1
C 37	73.2	7.3	14859	12	US-10-051-874-55	Sequence 55, Appl
C 38	72.6	7.2	1388	12	US-10-264-049-518	Sequence 518, App
C 39	72.2	7.2	469	13	US-10-029-386-20613	Sequence 20613, A
C 40	72.2	7.2	5379	15	US-10-128-714-5305	Sequence 5305, Ap
C 41	71.8	7.1	2404	12	US-10-104-047-1740	Sequence 1740, Ap
C 42	71.8	7.1	3439	12	US-10-161-927-53	Sequence 53, Appl
C 43	71.6	7.1	628	13	US-10-029-386-22859	Sequence 22859, A
C 44	71.6	7.1	1008	9	US-09-922-480-1	Sequence 1, Appl1
C 45	71.6	7.1	1008	9	US-09-923-236-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-09-922-480-1  
Sequence 1, Application US/09922480  
Patent No. US20020081701A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Adler, David A.  
TITLE OF INVENTION: SECRETED SALIVARY ZS1G63 POLYPEPTIDE  
FILE REFERENCE: 97-71  
CURRENT APPLICATION NUMBER: US/09/922,480  
PRIOR FILING DATE: 2001-08-03  
PRIOR APPLICATION NUMBER: US 60/124,820  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 1008  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (128)...(784)  
US-09-922-480-1

Query Match 100.0%; Score 1008; DB 9; Length 1008;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACAGACCTAAAGGCGATGTTCTTCCTCTAAAGAGAAATATTAATT 60  
DB 1 AGACAGACCTAAAGGCGATGTTCTTCCTCTAAAGAGAAATATTAATT 60  
QY TAAATATGATTCGATTTCTTAAACATTAATTATAGTTATATTCAGGTC 120  
DB TAAATATGATTCGATTTCTTAAACATTAATTATAGTTATATTCAGGTC 120  
QY AATCAATGAAAGCTTCCTTGGGCTGCATGTATGTGCTTTGCAAGAGAG 180

```

Db 121 AATCAAAATGAAGCTTCTCTTGGGCGCTGCATGTGATGTGCTTTTGCAAGGAAAG 180
Qy 181 ACGGTTCCCTTCATTTGGAGAGATGACAATGAGATGTCACCCACTTCAATCTCT 240
Db 181 ACGGTTCCCTTCATTTGGAGAGATGACAATGAGATGTCACCCACTTCAATCTCT 240
Qy 241 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAAATAC 300
Db 241 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAAATAC 300
Qy 301 AGTCCCAAGTTACCTTGGAAATCTTAACATGACACAGGGTTACCTTGATTCCTTGAT 360
Db 301 AGTCCCAAGTTACCTTGGAAATCTTAACATGACACAGGGTTACCTTGATTCCTTGAT 360
Qy 361 TCTAATCTTCTCTGATTCCTTATGTCTATCAATCCGTTGTTCCCTTACTACTCA 420
Db 361 TCTAATCTTCTCTGATTCCTTATGTCTATCAATCCGTTGTTCCCTTACTACTCA 420
Qy 421 GTTGAATGTTCTCTCTCTCTCTCTAGGGGTTTCCGTTTGTCTCTCTCAAGGTTT 480
Db 421 GTTGAATGTTCTCTCTCTCTCTCTAGGGGTTTCCGTTTGTCTCTCTCAAGGTTT 480
Qy 481 TTGACAGCTGACAGACCCGCTGCCACCTATTCAGAGCTGAGCTGCTGACCTGAC 540
Db 481 TTGACAGCTGACAGACCCGCTGCCACCTATTCAGAGCTGAGCTGCTGACCTGAC 540
Qy 541 TCTTACAGCCACCTGTAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 600
Db 541 TCTTACAGCCACCTGTAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 600
Qy 601 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 660
Db 601 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 660
Qy 661 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 720
Db 661 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 720
Qy 721 TGCTACAGCCAGCTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 780
Db 721 TGCTACAGCCAGCTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 780
Qy 781 TCAAGTGAATTCCTAGAGAGTACATGAGGTTCTATTCATGATGACAAATTAAGT 840
Db 781 TCAAGTGAATTCCTAGAGAGTACATGAGGTTCTATTCATGATGACAAATTAAGT 840
Qy 841 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAAATAC 900
Db 841 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAAATAC 900
Qy 901 ATTATCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901 ATTATCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 961 AAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
Db 961 AAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008

```

```

RESULT 2
US-09-923-236-1
; Sequence 1, Application US/09923236
; Patent No. US20020090677A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/923,236
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 9

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128) ... (784)
US-09-923-236-1

```

```

Query Match 100.0%; Score 1008; DB 9; Length 1008;
Best Local Similarity 100.0%; Pred. No. 1..e-236; Indels 0; Gaps 0;
Matches 1008; Conservative 0; Mismatches 0;

```

```

Qy 1 AGACAGACTAAAGGACATGATTCCTTGTCTCTCTTAAAGAGAAATATTAAT 60
Db 1 AGACAGACTAAAGGACATGATTCCTTGTCTCTCTTAAAGAGAAATATTAAT 60
Qy 61 TAAATATACATGGGATTTCTTAAACAAATTAATAGTTAATTCATAGGTC 120
Db 61 TAAATATACATGGGATTTCTTAAACAAATTAATAGTTAATTCATAGGTC 120
Qy 121 AATCAAAATGAAGCTTCTCTTGGGCTGATGTATGTGTCTTTTGCAGAGAG 180
Db 121 AATCAAAATGAAGCTTCTCTTGGGCTGATGTATGTGTCTTTTGCAGAGAG 180
Qy 181 ACGGTTCCCTTCATTTGGAGAGATGACAATGAGATGTCACCCACTTCAATCTCT 240
Db 181 ACGGTTCCCTTCATTTGGAGAGATGACAATGAGATGTCACCCACTTCAATCTCT 240
Qy 241 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAA 300
Db 241 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAA 300
Qy 301 AGTCCCAAGTTACCTTGGAAATCTTAACATGACACAGGGTTACCTTGATTCCTTGAT 360
Db 301 AGTCCCAAGTTACCTTGGAAATCTTAACATGACACAGGGTTACCTTGATTCCTTGAT 360
Qy 361 TCTAATCTTCTCTGATTCCTTATGTCTATCAATCCGTTGTTCCCTTACTACTCA 420
Db 361 TCTAATCTTCTCTGATTCCTTATGTCTATCAATCCGTTGTTCCCTTACTACTCA 420
Qy 421 GTTGAATGTTCTCTCTCTCTCTCTAGGGGTTTCCGTTTGTCTCTCTCAAGGTTT 480
Db 421 GTTGAATGTTCTCTCTCTCTCTCTAGGGGTTTCCGTTTGTCTCTCTCAAGGTTT 480
Qy 481 TTGACAGCTGACAGACCCGCTGCCACCTATTCAGAGCTGAGCTGCTGACCTGAC 540
Db 481 TTGACAGCTGACAGACCCGCTGCCACCTATTCAGAGCTGAGCTGCTGACCTGAC 540
Qy 541 TCTTACAGCCACCTGTAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 600
Db 541 TCTTACAGCCACCTGTAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 600
Qy 601 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 660
Db 601 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 660
Qy 661 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 720
Db 661 TGCTGACAGAGGCACTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 720
Qy 721 TGCTACAGCCAGCTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 780
Db 721 TGCTACAGCCAGCTGTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 780
Qy 781 TCAAGTGAATTCCTAGAGAGTACATGAGGTTCTATTCATGATGACAAATTAAGT 840
Db 781 TCAAGTGAATTCCTAGAGAGTACATGAGGTTCTATTCATGATGACAAATTAAGT 840
Qy 841 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAAATAC 900
Db 841 GAATATTCCTTAATGAGATGAGAAATTTACCACTTCTTTATTAATGAGGAGAAATAC 900

```





```

QY 672 CACCTGTTGAGTGAGCCAGCTGACAGAGAACTTCAACAGCTGAGCTGCTACAGCA 731
DB 612 CACCTGTTGAGTGAGCCAGCTGACAGAGAACTTCAACAGCTGAGCTGCTACAGCA 671
QY 732 AGCTGCTGCCCCAGAACTCAACCTCTTCCCTCTCTTGAACAGGCAATCAGTGAATT 791
DB 672 AGCTGCTGCCCCAGAACTCAACCTCTTCCCTCTCTTGAACAGGCAATCAGTGAATT 731
QY 792 CTCCTGAAGAGTACCAATGGCTTCACTTCTACTGATGAGAAATAAGTGAATCTACAA 851
DB 732 CTCCTGAAGAGTACCAATGGCTTCACTTCTACTGATGAGAAATAAGTGAATCTACAA 791
QY 852 AAGTTTCTTCTTCTTCCAAAGACTATTTCACTCTGTGATTCAGAGATTCACTCAC 911
DB 792 AAGTTTCTTCTTCTTCCAAAGACTATTTCACTCTGTGATTCAGAGATTCACTCAC 851
QY 912 TACATTGATTTGTTGTGTGAGTTTCTTGTGACTTAAATTAATTAAGAAAAACATTGA 971
DB 852 TACATTGATTTGTTGTGTGAGTTTCTTGTGACTTAAATTAATTAAGAAAAACATTGA 911
QY 972 TAATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1005
DB 912 TAATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 945

```

## RESULT 6

```

US-09-922-480-3
/ Sequence 3, Application US/09922480
/ Patent No. US20020081701A1
/ GENERAL INFORMATION:
/ APPLICANT: Sheppard, Paul O.
/ APPLICANT: Adler, David A.
/ TITLE OF INVENTION: SECRETED SALIVARY ZS1G63 POLYPEPTIDE
/ FILE REFERENCE: 97-71
/ CURRENT APPLICATION NUMBER: US/09/922,480
/ CURRENT FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: US 60/124,820
/ PRIOR FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3
/ LENGTH: 657
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Degenerate polynucleotide sequence for zsig63
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(657)
/ OTHER INFORMATION: n = A,T,C or G
US-09-922-480-3

```

```

Query Match 45.6%; Score 460; DB 9; Length 657;
Best Local Similarity 61.0%; Pred. No. 1.9e-102;
Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

```

```

QY 128 ATGAAGCTTCTCTTGGGCGCTGATGTATGTGCTTTTGCAGGAAGAGCGCTTC 187
DB 1 ATGAAGCTTCTCTTGGGCGCTGATGTATGTGCTTTTGCAGGAAGAGCGCTTC 60
QY 188 CCCTTCATTTGAGAGATGACATGACGATGTCAACCACTTCATCATCTCTGATATT 247
DB 61 CCCTTCATTTGAGAGATGACATGACGATGTCAACCACTTCATCATCTCTGATATT 120
QY 248 CCTTATGCAATACGGAATTACCACTCTCTTATATATGCGCCAGTGAATACAGTCCC 307
DB 121 CCTTATGCAATACGGAATTACCACTCTCTTATATATGCGCCAGTGAATACAGTCCC 180
QY 308 AGTTACCTGAGGAATCTTACACTGACACAGGGTTACCTGATCCCTGAGATTCTAAT 367
DB 181 AGTTACCTGAGGAATCTTACACTGACACAGGGTTACCTGATCCCTGAGATTCTAAT 240
QY 368 TCTCTGAGATTCCCTATGTCTATCAATCCGTTGTTTCCCTTACCTACCTGATTAAT 427

```

```

DB 241 WSNCCNGNNTTTCNTATAGTATAYCAVATHMNGNNTTTCNTATGACNACATYNAAY 300
QY 428 GTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487
DB 301 GTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 488 GCTGAGACCCGCTGCCCCCACTTATGAGCTGAGCTGCTGAGCTGACCTCTTACA 547
DB 361 GCTGAGACCCGCTGCCCCCACTTATGAGCTGAGCTGCTGAGCTGACCTCTTACA 420
QY 548 GCGACACTGTAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 607
DB 421 GCGACACTGTAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 480
QY 608 GAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 667
DB 481 GAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 540
QY 668 GAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 727
DB 541 GAGGACCTGTTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 600
QY 728 GCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784
DB 601 GCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657

```

## RESULT 7

```

US-09-923-236-3
/ Sequence 3, Application US/09923236
/ Patent No. US20020090677A1
/ GENERAL INFORMATION:
/ APPLICANT: Sheppard, Paul O.
/ APPLICANT: Adler, David A.
/ TITLE OF INVENTION: SECRETED SALIVARY ZS1G63 POLYPEPTIDE
/ FILE REFERENCE: 97-71
/ CURRENT APPLICATION NUMBER: US/09/923,236
/ CURRENT FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: US 60/124,820
/ PRIOR FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3
/ LENGTH: 657
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Degenerate polynucleotide sequence for zsig63
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(657)
/ OTHER INFORMATION: n = A,T,C or G
US-09-923-236-3

```

```

Query Match 45.6%; Score 460; DB 9; Length 657;
Best Local Similarity 61.0%; Pred. No. 1.9e-102;
Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

```

```

QY 128 ATGAAGCTTCTCTTGGGCGCTGATGTATGTGCTTTTGCAGGAAGAGAGCGCTTC 187
DB 1 ATGAAGCTTCTCTTGGGCGCTGATGTATGTGCTTTTGCAGGAAGAGAGCGCTTC 60
QY 188 CCCTTCATTTGAGAGATGACATGACGATGTCAACCACTTCATCATCTCTGATATT 247
DB 61 CCCTTCATTTGAGAGATGACATGACGATGTCAACCACTTCATCATCTCTGATATT 120
QY 248 CCTTATGCAATACGGAATTACCACTCTCTTATATATGCGCCAGTGAATACAGTCCC 307
DB 121 CCTTATGCAATACGGAATTACCACTCTCTTATATATGCGCCAGTGAATACAGTCCC 180
QY 308 AGTTACCTGAGGAATCTTACACTGACACAGGGTTACCTGATCCCTGAGATTCTAAT 367
DB 181 AGTTACCTGAGGAATCTTACACTGACACAGGGTTACCTGATCCCTGAGATTCTAAT 240

```

QY 368 TCTCTGATTCCTCTATGTATCATCCGCTGTTTCCCTTAGTACTAGTTGAAT 427  
DB 241 MSNCCNGANTTYCCNTAAGTNTAYCAATMNGNGNTTYCCNYTNGCNAACARNTNAY 300  
QY 428 GTTCTCTCTCCCTCTAGAGGTTTCCGTTTGTCCCTCTTCAAGTTTTCACAGA 487  
DB 301 GTNCCNCCNYTNCNCCMNGNGNTTYCCNTTYGTNCCNCCMSMNGNTTYTTTWSNCGN 360  
QY 488 GCTGAGACCCGCTGCCCCACCTATTGAGCTGAGCTGTCAGCTGACCTCTTACA 547  
DB 361 GGNCGNCCNGCNGCNGCNCNCCNATHGNGCNGARCCNGCNGCNGCNCNCTNACN 420  
QY 548 GCCACACCTGTAGACAGCTGAGCTGTGAGAGGCCCCCTGTGTGACAGCTGCTGCA 607  
DB 421 GGNACNCCNGTNGCNGCNGARCCNCGNCGNCGNCCNGTNGCNGCNGARCCNCGN 480  
QY 608 GAGGACCTGTGTGAGCTGAGCTGCTGTCAGAGGACCTGTGTGACAGCTGCTGCA 667  
DB 481 GARGCNCNGTNGCNGCNGARCCNCGNCGNCGNCGNCCNGTNGCNGCNGARCCNCGN 540  
QY 668 GAGGACCTGTGTGAGCTGAGCTGTCAGAGGAACTTCAACAGCTGAGCTGCTACA 727  
DB 541 GARGCNCNGTNGCNGCNGARCCNCGNCGNCGNCGNCGNCCNMSNCCNGCNGARCCNCGN 600  
QY 728 GCCAAGCTGTGCTGCCCGGACCACTCACTCTCTCTCTCTTGAACAGGCAATCAG 784  
DB 601 GGNARCCNCGNCCNGCNGARCCNCCNCCNMSNCCNMSNTNGARCCAGCNAAYCAR 657

# RESULT 8

US-09-922-469-3

/ Sequence 3, Application US/09922469  
/ Patent No. US20020173027A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Shepherd, Paul O.  
/ APPLICANT: Adler, David A.  
/ TITLE OF INVENTION: SECRETED SALLIVARY ZS1G63 POLYPEPTIDE  
/ FILE REFERENCE: 97-71  
/ CURRENT APPLICATION NUMBER: US/09/922,469  
/ CURRENT FILING DATE: 2001-08-03  
/ PRIOR APPLICATION NUMBER: US 60/124,820  
/ PRIOR FILING DATE: 1999-03-17  
/ NUMBER OF SEQ ID NOS: 9  
/ SOFTWARE: FastSeq for Windows Version 3.0  
/ SEQ ID NO 3  
/ LENGTH: 657  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Degenerate polynucleotide sequence for zs1g63  
/ NAME/KEY: misc\_feature  
/ LOCATION: (1)...(657)  
/ OTHER INFORMATION: n = A,T,C or G  
US-09-922-469-3

Query Match 45.6%; Score 460; DB 10; Length 657;  
Best Local Similarity 61.0%; Pred. No. 1.9e-102;  
Matches 401; Conservative 103; Mismatches 153; Indels 0; Gaps 0;

QY 128 ATGAAGCTTCTCTTGGGCTGCAATGTATGTGTGCTTTGCAAGAGAGAGGCTC 187  
DB 1 ATGAATTTTNTTNTGGGCTGTATGTGTGTGCTTTGCAAGAGAGAGGNTTY 60  
QY 188 CCGTTCATTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 247  
DB 61 CCGTTCATTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
QY 248 CCGTTCATTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 307  
DB 121 CCGTTCATTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 308 AGTACCTCTGGAATACTTACATGACAGAGGTTACTTCTGATCCCTGATCTTAAT 367

DB 181 MSNTAYCCNGAANAATCTATACAGATACNGANTTNCNMSNTATCCATGATHTYTNACN 240  
QY 368 TCTCTGATTCCTCTATGTATCATCCGCTGTTTCCCTTAGTACTAGTTGAAT 427  
DB 241 MSNCCNGANTTYCCNTAAGTNTAYCAATMNGNGNTTYCCNYTNGCNAACARNTNAY 300  
QY 428 GTTCTCTCTCCCTCTAGAGGTTTCCGTTTGTCCCTCTTCAAGTTTTCACAGA 487  
DB 301 GTNCCNCCNYTNCNCCMNGNGNTTYCCNTTYGTNCCNCCMSMNGNTTYTTTWSNCGN 360  
QY 488 GCTGAGACCCGCTGCCCCACCTATTGAGCTGAGCTGTCAGCTGACCTCTTACA 547  
DB 361 GGNCGNCCNGCNGCNGCNCNCCNATHGNGCNGARCCNGCNGCNGCNCNCTNACN 420  
QY 548 GCCACACCTGTAGACAGCTGAGCTGTGAGAGGCCCCCTGTGTGACAGCTGCTGCA 607  
DB 421 GGNACNCCNGTNGCNGCNGARCCNCGNCGNCGNCCNGTNGCNGCNGARCCNCGN 480  
QY 608 GAGGACCTGTGTGAGCTGAGCTGCTGTCAGAGGACCTGTGTGACAGCTGCTGCA 667  
DB 481 GARGCNCNGTNGCNGCNGARCCNCGNCGNCGNCGNCCNGTNGCNGCNGARCCNCGN 540  
QY 668 GAGGACCTGTGTGAGCTGAGCTGTCAGAGGAACTTCAACAGCTGAGCTGCTACA 727  
DB 541 GARGCNCNGTNGCNGCNGARCCNCGNCGNCGNCGNCGNCCNMSNCCNGCNGARCCNCGN 600  
QY 728 GCCAAGCTGTGCTGCCCGGACCACTCACTCTCTCTCTCTTGAACAGGCAATCAG 784  
DB 601 GGNARCCNCGNCCNGCNGARCCNCCNCCNMSNCCNMSNTNGARCCAGCNAAYCAR 657

# RESULT 9

US-10-294-804-1/c

/ Sequence 1, Application US/10294804  
/ Publication No. US20030133948A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Robertson, Eric S.  
/ APPLICANT: Cotter, Murray A.  
/ TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA  
/ FILE REFERENCE: UM-03778  
/ CURRENT APPLICATION NUMBER: US/10/294,804  
/ CURRENT FILING DATE: 2002-11-14  
/ PRIOR APPLICATION NUMBER: US/09/410,399  
/ PRIOR FILING DATE: 1999-10-01  
/ NUMBER OF SEQ ID NOS: 6  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 1  
/ LENGTH: 3489  
/ TYPE: DNA  
/ ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-10-294-804-1

Query Match 12.0%; Score 121.2; DB 13; Length 3489;  
Best Local Similarity 58.7%; Pred. No. 6.5e-19;  
Matches 210; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 413 GCTACTGATGTAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 472  
DB 2188 GCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 2129  
QY 413 AGCTTTTATGAGACAGCTGAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069  
DB 2128 CCGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069  
QY 533 GCTGACCTCTTTAGAGCACAACCTGTAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 592  
DB 2068 GCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009  
QY 593 GCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652  
DB 2008 GCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 07:13:18 ; Search time 2130 Seconds  
(without alignments)  
11501.837 Million cell updates/sec

Title: US-09-923-236-1

Perfect score: 1008  
Sequence: 1 agacagacataaaaaagccat.....tagataattagaccaatg 1008

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pug:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	703	69.7	703	2	BSM075873
2	680.4	67.5	742	14	BSM075873
3	677.4	67.2	742	14	BSM075873
4	667	66.2	677	2	BSM075890

Result No.	Score	Query Match	Length	DB ID	Description
5	655	65.0	655	2	BSM075629
6	655	65.0	655	2	BSM075629
7	654.6	64.9	654.6	2	BSM075629
8	647	64.2	647	2	BSM075629
9	641.4	63.6	641.4	2	BSM075629
10	640.4	63.5	640.4	2	BSM075629
11	633.4	62.8	633.4	2	BSM075629
12	614.2	60.9	614.2	2	BSM075629
13	605	60.0	605	2	BSM075629
14	603.4	59.9	603.4	2	BSM075629
15	585.4	58.1	585.4	2	BSM075629
16	562	55.8	562	2	BSM075629
17	549.4	54.5	549.4	2	BSM075629
18	541.8	53.8	541.8	2	BSM075629
19	519	51.5	519	2	BSM075629
20	516.8	51.3	516.8	2	BSM075629
21	490.6	48.7	490.6	2	BSM075629
22	483.8	48.0	483.8	2	BSM075629
23	483.8	48.0	483.8	2	BSM075629
24	476	47.2	476	2	BSM075629
25	467.8	46.4	467.8	2	BSM075629
26	451	44.7	451	2	BSM075629
27	433.6	43.0	433.6	2	BSM075629
28	403.8	40.1	403.8	2	BSM075629
29	394	39.1	394	2	BSM075629
30	299	29.7	299	2	BSM075629
31	258.2	25.6	258.2	2	BSM075629
32	227	22.5	227	2	BSM075629
33	215	21.3	215	2	BSM075629
34	178.2	17.7	178.2	2	BSM075629
35	153	15.2	153	2	BSM075629
36	133	13.2	133	2	BSM075629
37	131.2	13.0	131.2	2	BSM075629
38	131	12.5	131	2	BSM075629
39	125.8	12.5	125.8	2	BSM075629
40	109.6	10.9	109.6	2	BSM075629
41	109.6	10.9	109.6	2	BSM075629
42	109.6	10.9	109.6	2	BSM075629
43	108	10.7	108	2	BSM075629
44	108	10.7	108	2	BSM075629
45	108	10.7	108	2	BSM075629

## ALIGNMENTS

RESULT 1  
ID HSM075873 standard; RNA; EST; 703 BP.  
XX HSM075873  
AC BX485686.1  
SV BX485686.1  
DT 09-MAY-2003 (Rel. 75, Created)  
DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)  
XX Homo sapiens mRNA; EST DKFZp686C15248\_r1 (from clone DKFZp686C15248)  
XX EST; expressed sequence tag.  
XX Homo sapiens (human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
XX (1)  
XX 1-703  
XX Octenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.W., Well B.,  
XX Amid C., Oeanger A., Fobo G., Han M., Wiemann S.,  
XX Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
XX MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY



```

Db      181 ATGGCATACGGAATTACCACTCTCTTATTATGCCCAGTAAGTAAGTCCCAAGT 240
Qy      312 ACCCTGGGAATATTACCTGACAGACAGGGTACTCTGATCCCTGGATTTCTAACTTTC 371
Db      241 ACCCTGGGAATATTACCTGACAGACAGGGTACTCTGATCCCTGGATTTCTAACTTTC 300
Qy      372 CTGGATTTCCCTATGCTATCATCATCCGAGTGTTCCTTAAAGTACTGAGTTGATGTC 431
Db      301 CTGGATTTCCCTATGCTATCATCATCTGAGTGTTCCTTAAAGTACTGAGTTGATGTC 360
Qy      432 CTCTCTCTCTCTCTAGAGGGGTTTCCGTTTGTCCCTCTCTCAAGTTTCTTTCAGAGCTG 491
Db      361 CTCTCTCTCTCTCTAGAGGGGTTTCCGTTTGTCCCTCTCTCAAGTTTCTTTCAGAGCTG 420
Qy      492 CAGCACCCTGCTGCCCACTATGACAGCTGAGGCTCTGAGAGTGAAGCTTTTACAGCCA 551
Db      421 CAGCACCCTGCTGCCCACTATGACAGCTGAGGCTCTGAGAGTGAAGCTTTTACAGCCA 480
Qy      552 CACCTGTAGCAGCTGAGGCTGTGACAGAGGCTCTGTGACAGCTGAGGCTGTGACAGAG 611
Db      481 CACCTGTAGCAGCTGAGGCTGTGACAGAGGCTCTGTGACAGCTGAGGCTGTGACAGAG 540
Qy      612 CACCTGTGAGCTGAGGCTGTGACAGAGGCTGTGACAGCTGAGGCTGTGACAGAG 671
Db      541 CACCTGTGAGCTGAGGCTGTGACAGAGGCTGTGACAGCTGAGGCTGTGACAGAG 600
Qy      672 CACCTGTGAGCTGAGGCTGTGACAGAGGCTGTGACAGCTGAGGCTGTGACAGAG 731
Db      601 CACCTGTGAGCTGAGGCTGTGACAGAGGCTGTGACAGCTGAGGCTGTGACAGAG 660
Qy      732 AGGCTGTGAGGCTGTGACAGAGGCTGTGACAGGCTGTGACAGGCTGTGACAGGCT 753
Db      661 AGGCTGTGAGGCTGTGACAGAGGCTGTGACAGGCTGTGACAGGCTGTGACAGGCT 682

```

```

RESULT 3
CB957562      742 bp      mRNA      linear      EST 29-APR-2003
LOCUS      CB957562
DEFINITION      AGENCOURT 13785505 NIH_MGC_184 Homo sapiens cDNA clone
IMAGE:30352440 5', mRNA sequence.
ACCESSION      CB957562
VERSION      CB957562.1 GI:30213679
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDCM151 row: j column: 01
High quality sequence stop: 551.
Location/Qualifiers
1.742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30352440"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_184"
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site_1:
SfiI (ggccatcggc); Site_2: SfiI (ggccgctcggc);
Library is oligo-dT primed and directionally cloned. cDNA

```

was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGACATTAAGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCGGAGGCGGAGGAGTGG-dt(30)-BN-3' (where B = A, C, G or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT 172 a 203 c 154 g 213 t

Query Match 67.2%; Score 677.4; DB 14; Length 742;  
Best Local Similarity 98.8%; Pred. No. 1.5e-99;  
Matches 714; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

```

Qy      1 AGACAGACTAAAAAGCCATGATCTTCTCTCTTAAAGAGAAAAATATAT 60
Db      16 AGACAGACTAAAAAGCCATGATCTTCTCTCTTAAAGAGAAAAATATAT 75
Qy      61 TAAAAATACATTCGATTTTCTAAACATTAATTTATAGTTATATTCATAGGCTC 120
Db      76 TAAAAATACATTCGATTTTCTAAACATTAATTTATAGTTATATTCATAGGCTC 135
Qy      121 AATCAAAATGAAGCTTCTCTTGGGCGGCAATGATATGTTGCTTTGCAAGAGAG 180
Db      136 AATCAAAATGAAGCTTCTCTTGGGCGGCAATGATATGTTGCTTTGCAAGAGAG 195
Qy      181 ACGGTTCCCTCATTTGGTGAAGATGACATAGCATGATGTCACCACTTCATCATCTCT 240
Db      196 ACGGTTCCCTCATTTGGTGAAGATGACATAGCATGATGTCACCACTTCATCATCTCT 255
Qy      241 GAATATTCCTTATGCAATACGAAATTTACACCTCTCTTATATATGCGCCAGTGAATAC 300
Db      256 GAATATTCCTTATGCAATACGAAATTTACACCTCTCTTATATATGCGCCAGTGAATAC 315
Qy      301 AGTCCCAAGTTACCTGGGAATACCTTACATGACACAGGGTATACCTGGTATCCCGGAT 360
Db      316 AGTCCCAAGTTACCTGGGAATACCTTACATGACACAGGGTATACCTGGTATCCCGGAT 375
Qy      361 TCTAATCTCTCTGATTCCTCTATGTCATACATCCGATGTTTCCCTTATGACTATCA 420
Db      376 TCTAATCTCTCTGATTCCTCTATGTCATACATCCGATGTTTCCCTTATGACTATCA 435
Qy      421 GTTGAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db      436 GTTGAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 495
Qy      481 TTACAGAGCTGACAGACCGGCTGCCCACTTATGACGCTGAGGCTGAGCTGACACC 540
Db      496 TTACAGAGCTGACAGACCGGCTGCCCACTTATGACGCTGAGGCTGAGCTGACACC 555
Qy      541 TCTTACAGCACAACCTGTAGCAGCTGAGCTGAGGAGGAGGCTGTGAGCTGAGGAGCC 600
Db      556 TCTTACAGCACAACCTGTAGCAGCTGAGCTGAGGAGGAGGCTGTGAGCTGAGGAGCC 615
Qy      601 TCTTACAGCACAACCTGTGAG-AGCTGAGCTGCTGACAGAGGACCTGTGAGCTGAGC 659
Db      616 TCTTACAGCACAACCTGTGAGAGCTGAGGAGGCTGACAGAGCACTGTGAGCTGAGC 675
Qy      660 CTGCTGAGAGGAGCAGCTG-TTGAGTGAAGCAGCTGAGAGAGAA-CCTTACAGAGCTGA 717
Db      676 CTGCTGAGAGAGCAGCTGTTTGAATGAATCCGCTGAGAGAAACCTTTCACAGCTGA 735
Qy      718 GCC 720
Db      736 GCC 738

```

RESULT 4  
HSW075890 standard; RNA; EST; 677 BP.  
ID HSW075890







QY 550 CACACCTGTAGACGTGAGCTGTCAGAGGGCCCTGTTGAGCTGAGCCTGTCAGCA 609  
 Db 563 CACACCTGTAGACGTGAGCTGTCAGAGGGCCCTGTTGAGCTGAGCCTGTCAGCA 622  
 QY 610 GGCACCTGTGAGCTGAGCCTGTCAGAGGACCTGTTGAGCTGAGCCTGCT 664  
 Db 623 GGCACCTGTGAGCTGAGCCTGTCAGAGGACCTGTTGAGCTGAGCCTGCT 677

## RESULT 7

HSN076755  
 ID HSN076755 standard; RNA; EST; 675 BP.

XX BX486568;  
 XX BX486568.1  
 SV BX486568.1  
 XX  
 DT 09-MAY-2003 (Rel. 75, Created)  
 DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)  
 XX  
 DE Homo sapiens mRNA; EST DKFZp686H08252\_r1 (from clone DKFZp686H08252)  
 XX  
 XX EST; expressed sequence tag.

OS Homo sapiens (human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homidae; Homo.

XX [1]  
 RN 1-675  
 RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,  
 RA Han M., Wiemann S.;  
 RL Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX This is the 5' sequence of the clone insert  
 CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 CC sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
 CC consortium of the German Genome Project.

CC No sl sequence available.  
 CC This clone (DKFZp686H08252) is available at the RZPD in Berlin.  
 CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,  
 CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de

XX Key Location/Qualifiers

FT source 1. 675  
 FT /db\_xref="taxon:9606"  
 FT /mol\_type="RNA"  
 FT /organism="Homo sapiens"  
 FT /clone="DKFZp686H08252"  
 FT /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
 FT DH10B; sites SfiI + SfiIb"  
 FT /dev\_stage="adult"  
 FT /cissue\_type="cDNA-collection"

SQ Sequence 675 BP; 139 A; 194 C; 150 G; 191 T; 1 other;

Query Match 64.9%; Score 654.6; DB 2; Length 675;  
 Best Local Similarity 99.3%; Pred. No. 7e-96;  
 Matches 668; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 68 ACATTCGATTTTCTAAACATAAATTAATGTTAATATTCAT-AGGGTCATGCA 126  
 Db 3 ACATTCGATTTTCTAAACATAAATTAATGTTAATATTCAT-AGGGTCATGCA 62  
 QY 127 AATGAAGCTTCTCTTGGGCTGCAATGTATGTGTTTGGCAAGAGAGACGTT 186  
 Db 63 AATGAAGCTTCTCTTGGGCTGCAATGTATGTGTTTGGCAAGAGAGACGTT 122  
 QY 187 CCCCTTCATTGGAGAGATGCAATGATGTCACCCACTTCATCTCTGAAATAT 246

Db 123 CCCCTTCATTGGAGATGCAATGATGTCACCCACTTCATCTCTGAAATAT 182  
 QY 247 TCCTTAAGCATAGGAATTTTACACCTCCCTTTAATATGCGCAAGTAATACAGTCC 306  
 Db 183 TCCTTAAGCATAGGAATTTTACACCTCCCTTTAATATGCGCAAGTAATACAGTCC 242  
 QY 307 CAGTTACCTCGGAATATTACACTGACACAGGGTTACCTTGATCCCTGAAATCTAAC 366  
 Db 243 CAGTTACCTCGGAATATTACACTGACACAGGGTTACCTTGATCCCTGAAATCTAAC 302  
 QY 367 TTCTCCGGAATTTCCCTTAATGCTATCAATCCGTTGTTTCCCTTAAGTACTAGTGA 426  
 Db 303 TTCTCCGGAATTTCCCTTAATGCTATCAATCCGTTGTTTCCCTTAAGTACTAGTGA 362  
 QY 427 TGTTCCTCCTCCCTCTAGAGGTTTCCGTTTCCCTCTTCAAGTTTTCAGC 486  
 Db 363 TGTTCCTCCTCCCTCTAGAGGTTTCCGTTTCCCTCTTCAAGTTTTCAGC 422  
 QY 487 AGCTGACAGACCCGCTGCCCCCACTATTGACAGTGAAGCTGACGCTGCACTCTTAC 546  
 Db 423 AGCTGACAGACCCGCTGCCCCCACTATTGACAGTGAAGCTGACGCTGCACTCTTAC 482  
 QY 547 AGCCACACCTGTAGACGTGAGCTGTCAGAGGGCCCTGTTGAGCTGAGCCTGCTGC 606  
 Db 483 AGCCACACCTGTAGACGTGAGCTGTCAGAGGGCCCTGTTGAGCTGAGCCTGCTGC 542  
 QY 607 AGAGCACCTGTGAGCTGAGCTGTCAGAGGACCTGTTGAGCTGAGCCTGCTGC 666  
 Db 543 AGAGCACCTGTGAGCTGAGCTGTCAGAGGACCTGTTGAGCTGAGCCTGCTGC 602  
 QY 667 AGAGCACCTGTGAGCTGAGCTGTCAGAGGACCTGTTGAGCTGAGCCTGCTGC 726  
 Db 603 AGAGCACCTGTGAGCTGAGCTGTCAGAGGACCTGTTGAGCTGAGCCTGCTGC 662  
 QY 727 AGCCAGGCTGCT 739  
 Db 663 AGCCAGGCTGCT 675

## RESULT 8

HSN076748  
 ID HSN076748 standard; RNA; EST; 675 BP.

XX BX486561;  
 XX BX486561.1  
 SV BX486561.1

DT 09-MAY-2003 (Rel. 75, Created)  
 DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)  
 XX  
 DE Homo sapiens mRNA; EST DKFZp686H20252\_r1 (from clone DKFZp686H20252)  
 XX  
 XX EST; expressed sequence tag.

OS Homo sapiens (human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homidae; Homo.

XX [1]  
 RN 1-675  
 RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,  
 RA Han M., Wiemann S.;  
 RL Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX This is the 5' sequence of the clone insert  
 CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 CC sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
 CC consortium of the German Genome Project.  
 CC No sl sequence available.

CC This clone (DKFZp666H20252) is available at the RZPD in Berlin.  
 CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,  
 CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers

PH source

1. .675  
 /db\_xref="taxon:9606"  
 /mol\_type="mRNA"  
 /organism="Homo sapiens"  
 /clone="DKFZp666H20252"  
 /clone\_lib="686 (synonym: h1cc3). Vector pSport1\_Sfi; host  
 DH10B; sites SfiIA + SfiIB"  
 /dev\_stage="adult"  
 /tissue\_type="CDNA-collection"

XX Sequence 675 BP; 157 A; 183 C; 135 G; 200 T; 0 other;

Query Match 64.2%; Score 647; DB 2; Length 675;  
 Best Local Similarity 98.5%; Pred. No. 1.2e-94;

Matches 653; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 AGACAGACTAAAGGACATATCTTCTGCTTCTCTCTAAAGAGAAATATATAT 60
DB 13 AGACAGACTAAAGGACATATCTTCTGCTTCTCTCTAAAGAGAAATATATAT 72
QY 61 TAAATAATACATGCGTATTTCTTAAACAATAATTAATAGTATATTCATAGGCTC 120
DB 73 TAAATAATACATGCGTATTTCTTAAACAATAATTAATAGTATATTCATAGGCTC 132
QY 121 AATCAAAATGAAGCTTCTCTTTGGGCTGATTTGATGTTGCTTTGCAAGAAAG 180
DB 133 AATCAAAATGAAGCTTCTCTTTGGGCTGATTTGATGTTGCTTTGCAAGAAAG 192
QY 181 ACGGTTCCCTCTCTGATGAGATGACAAATGAGATGATCCCATCTTATCTCT 240
DB 193 ACGGTTCCCTCTCTGATGAGATGACAAATGAGATGATCCCATCTTATCTCT 252
QY 241 GAATATTCCTTATGAGATGAGATTTACCACTCTCTTATATATGCGCAAGTAATAC 300
DB 253 GAATATTCCTTATGAGATGAGATTTACCACTCTCTTATATATGCGCAAGTAATAC 312
QY 301 AGTCCCAAGTACCTGGAATACCTTACACTGACAGAGGTTACTTCTGATCCCTGAT 360
DB 313 AGTCCCAAGTACCTGGAATACCTTACACTGACAGAGGTTACTTCTGATCCCTGAT 372
QY 361 TCTAATCTTCTCTGATGATTTCCCTATGCTATCAATCCGTTGTTCCCTTATGCTATCA 420
DB 373 TCTAATCTTCTCTGATGATTTCCCTATGCTATCAATCCGTTGTTCCCTTATGCTATCA 432
QY 421 GTTGAATGTTCTCTCTCTCTCTCTGAGGTTTCCGTTGTTCCCTCTCAAGTTT 480
DB 433 GTTGAATGTTCTCTCTCTCTCTCTGAGGTTTCCGTTGTTCCCTCTCAAGTTT 492
QY 481 TTGAGAGCTGACAGACCCGCTGCTCCCACTTATGAGCTGAGCTGCTGACGCTGAC 540
DB 493 TTGAGAGCTGACAGACCCGCTGCTCCCACTTATGAGCTGAGCTGACGCTGAC 552
QY 541 TCTTACAGCCACACTTATGAGCTGAGCTGCTGCTGAGAGGCCCCCTTGTGAGCTGAGCC 600
DB 553 TCTTACAGCCACACTTATGAGCTGAGCTGCTGCTGAGAGGCCCCCTTGTGAGCTGAGCC 612
QY 601 TGCTGACAGAGCACTGTTGAGCTGAGCTGCTGACAGAGCACTTGTGAGCTGAGCC 660
DB 613 TGCTGACAGAGCACTGTTGAGCTGAGCTGCTGACAGAGCACTTGTGAGCTGAGCC 672
QY 661 TGC 663
DB 673 TGC 675

```

RESULT 9  
 HSM075590

ID HSM075590 standard; RNA; EST; 665 BP.

XX BX485403;

XX BX485403.1

SV 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX Homo sapiens mRNA; EST DKFZp686A05247\_x1 (from clone DKFZp686A05247)

XX EST; expressed sequence tag.

XX Homo sapiens (human)

XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

XX Rutheria; Primates; Catarrhini; Homnidae; Homo.

XX [1]

RP 1-665

RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Oeanger A., Fodo G.,

RA Han W., Wiemann S.;

RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPs, Ingolstaedter Landstr.1, D-85764 Neubertberg, GERMANY

XX This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

CC sequenced by Qiagen (Hilden/Germany) within the CDNA sequencing

CC consortium of the German Genome Project.

CC No 5' sequence available.

CC This clone (DKFZp686A05247) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers

PH source

1. .665  
 /db\_xref="taxon:9606"  
 /mol\_type="mRNA"  
 /organism="Homo sapiens"  
 /clone="DKFZp686A05247"  
 /clone\_lib="686 (synonym: h1cc3). Vector pSport1\_Sfi; host  
 DH10B; sites SfiIA + SfiIB"  
 /dev\_stage="adult"  
 /tissue\_type="CDNA-collection"

XX Sequence 665 BP; 152 A; 179 C; 133 G; 201 T; 0 other;

Query Match 63.6%; Score 641.4; DB 2; Length 665;  
 Best Local Similarity 99.7%; Pred. No. 9.5e-94;

Matches 653; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 AGACAGCTAAAGGACATGATTTCTTGGCTGATGATGTTGCTTTGCAAGAAAG 60
DB 12 AGACAGCTAAAGGACATGATTTCTTGGCTGATGATGTTGCTTTGCAAGAAAG 71
QY 61 TAAATAATACATGCGTATTTCTTAAACAATAATTAATAGTATATTCATAGGCTC 120
DB 72 TAAATAATACATGCGTATTTCTTAAACAATAATTAATAGTATATTCATAGGCTC 131
QY 121 AATCAAAATGAAGCTTCTCTTTGGGCTGATGATGTTGCTTTGCAAGAAAG 180
DB 132 AATCAAAATGAAGCTTCTCTTTGGGCTGATGATGTTGCTTTGCAAGAAAG 191
QY 181 ACGGTTCCCTCTTATGAGATGAGATGACAAATGAGATGATCCCATCTTCTCT 240
DB 192 ACGGTTCCCTCTTATGAGATGAGATGACAAATGAGATGATCCCATCTTCTCT 251
QY 241 GAATATTCCTTATGAGATGAGATTTACCACTCTCTTATATATGCGCAAGTAATAC 300
DB 252 GAATATTCCTTATGAGATGAGATTTACCACTCTCTTATATATGCGCAAGTAATAC 311

```

QY 301 AGTCCCGAGTTACCTGGGAATCTTACAGTACAGAGGTTACCTGGTATCCCTGGAT 360  
 Db 312 AGTCCCGAGTTACCTGGGAATCTTACAGTACAGAGGTTACCTGGTATCCCTGGAT 371  
 QY 361 TCTAATCTTCTGATGATTCCTTATGATGATGATGATGATGATGATGATGATGATGAT 420  
 Db 372 TCTAATCTTCTGATGATTCCTTATGATGATGATGATGATGATGATGATGATGATGAT 431  
 QY 421 GTTGAATGTCCT 480  
 Db 432 GTTGAATGTCCT 491  
 QY 481 TTCAAGAGTGAAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 Db 492 TTCAAGAGTGAAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551  
 QY 541 TCTTACAGCAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAG 600  
 Db 552 TCTTACAGCAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAG 611  
 QY 601 TGCTGAGAGGCACTGTTGAGAGTGAAGCTGCTGAGAGGCACTGTTGAGAGT 655  
 Db 612 TGCTGAGAGGCA-CTGTTGAGAGTGAAGCTGCTGAGAGGCACTGTTGAGAGT 665

## RESULT 10

HSM077932 standard; RNA; EST; 678 BP.

XX HSM077932  
 AC BX487745;  
 SV BX487745.1  
 DT 09-MAY-2003 (Rel. 75, Created)  
 DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

DE Homo sapiens mRNA; EST DKFZp686N16258\_r1 (from clone DKFZp686N16258)

XX EST; expressed sequence tag.

XX Homo sapiens (human)  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Euteria; Primates; Catarrhini; Homidae; Homo.

XX [1]  
 RN Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Robo G.,  
 RA Han M., Wiemann S.;

RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RL MIPs, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX This is the 5' sequence of the clone insert  
 CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 CC sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
 CC consortium of the German Genome Project.

CC No s1 sequence available.

CC This clone (DKFZp686N16258) is available at the RZPD in Berlin.

CC Please contact the RZPD, Ressourcenzentrum, Heubnerweg 6,  
 CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de

XX Key Location/Qualifiers

FT source 1..678  
 FT /db\_xref="taxon:9606"  
 FT /mol\_type="rRNA"  
 FT /organism="Homo sapiens"  
 FT /clone="DKFZp686N16258"  
 FT /clone\_1ib="686 (synonym: h1cc3). Vector pSPORT1\_sfi; host  
 FT DH10B; sites SfiI + SfiII  
 FT /dev\_stage="adult"  
 FT /issue\_type="cDNA-collection"

XX SQ Sequence 678 BP; 158 A; 180 C; 132 G; 207 T; 1 other;

Query Match 63.5%; Score 640.4; DB 2; Length 678;

Best Local Similarity 99.5%; Pred. No. 1.4e-93; Matches 652; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 AGACAGCTAATAAAGCATGATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
 Db 25 AGACAGCTAATAAAGCATGATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 84  
 QY 61 TAAATAATCATTTGGGATTTTCTTAAACATTAATTTATAGTTAATATCATAGGTC 120  
 Db 85 TAAATAATCATTTGGGATTTTCTTAAACATTAATTTATAGTTAATATCATAGGTC 144  
 QY 121 AATCAAAATGAAGTTCT 180  
 Db 145 AATCAAAATGAAGTTCT 204  
 QY 181 ACGGTTCCCTTCATTTGGTGAAGATGACATGACATGATGATGATGATGATGATGATGATGAT 240  
 Db 205 ACGGTTCCCTTCATTTGGTGAAGATGACATGACATGATGATGATGATGATGATGATGATGAT 264  
 QY 241 GAATATTCCTTATGAGCATAGGAAATTTACACCTCTCTCTTATATGAGCCAGTGAATAC 300  
 Db 265 GAATATTCCTTATGAGCATAGGAAATTTACACCTCTCTCTTATATGAGCCAGTGAATAC 324  
 QY 301 AGTCCCGAGTTACCTGGGAATCTTACAGTACAGAGGTTACCTGGTATCCCTGGAT 360  
 Db 325 AGTCCCGAGTTACCTGGGAATCTTACAGTACAGAGGTTACCTGGTATCCCTGGAT 384  
 QY 361 TCTAATCTTCTGATGATTCCTTATGATGATGATGATGATGATGATGATGATGATGAT 420  
 Db 385 TCTAATCTTCTGATGATTCCTTATGATGATGATGATGATGATGATGATGATGATGAT 444  
 QY 421 GTTGAATGTCCT 480  
 Db 445 GTTGAATGTCCT 504  
 QY 481 TTCAAGAGTGAAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 Db 505 TTCAAGAGTGAAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564  
 QY 541 TCTTACAGCAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAG 600  
 Db 565 TCTTACAGCAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAG 623  
 QY 601 TGCTGAGAGGCACTGTTGAGAGTGAAGCTGCTGAGAGGCACTGTTGAGAGT 655  
 Db 624 TGCTGAGAGGCACTGTTGAGAGTGAAGCTGCTGAGAGGCACTGTTGAGAGT 678

## RESULT 11

HSM077857 standard; RNA; EST; 647 BP.

XX HSM077857  
 AC BX487670;  
 SV BX487670.1

DT 09-MAY-2003 (Rel. 75, Created)  
 DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

DE Homo sapiens mRNA; EST DKFZp686G13258\_r1 (from clone DKFZp686G13258)

XX EST; expressed sequence tag.

XX Homo sapiens (human)  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Euteria; Primates; Catarrhini; Homidae; Homo.

XX [1]  
 RN 1-647

RA Bahr A., Lauber J., Mewes H.W., Well B., Amid C., Osanger A., Fobo G.,  
RA Han M., Wiemann S.;

Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases  
MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

CC This is the 5' sequence of the clone insert  
CC  
CC clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
CC sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
CC consortium of the German Genome Project.  
CC  
CC No. of sequences available: 1

CC This clone (DKFZp686G13258) is available at the RZPD in Berlin.  
CC Please contact the RZPD, [Rezeptionszentrum@rzpd.mpg.de](mailto:Rezeptionszentrum@rzpd.mpg.de)

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de  
XX

FH	Key	Location/Qualifiers
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

FT	Source	1. .647
FT	1. .647	

```
ET      /mol_type="mRNA"
ET      /organism="Homo sapiens"
```

```
ET /clone="DKFZp6866G13258"
```

```
F1      DHI0B; bites S11A + S11B"
ET      /dev stage="adult"
```

```

/ LIBRARY_type="CDNA-collection"
XX

```

[illegible]

Best Local Similarity 99.8%; Pred. No. 1.9e-92;

Matches 634; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AGACAGCTAAAAAGCATGTATTTCTTCGTTTCTCTAAABAAAGAAAAATTAATTT	60
Db	13	AGACAGCTAAAAAGCATGTATTTCTTCGTTTCTCTAAABAAAGAAAAATTAATTT	72
QY	61	TAAATAATACATTTGGCGTATTTTCTAAACAAATAAATTTATAGTAAATATTCATAGGGTC	120
Db	73	TAAATAATACATTTGGCGTATTTTCTAAACAAATAAATTTATAGTAAATATTCATAGGGTC	132
QY	121	AATCAAAATGAAGCTTCTCTTTGGGCTGCATGTATGTGTGTGCTTTTGCAAGGAAG	180
Db	133	AATCAAAATGAAGCTTCTCTTTGGGCTGCATGTATGTGTGTGCTTTTGCAAGGAAG	192
QY	181	ACGGTTCCTCTTCAATGGTAGAGATGACAAATGACATGATGACCACTTCACTCATCTCT	240
Db	193	ACGGTTCCTCTTCAATGGTAGAGATGACAAATGACATGATGACCACTTCACTCATCTCT	252
QY	241	GAATATTCCTTATGAGCATACGGAATTTACACACTCTCTTATATTAATGGCCAGTAATAC	300
Db	253	GAATATTCCTTATGAGCATACGGAATTTACACACTCTCTTATATTAATGGCCAGTAATAC	312
QY	301	AGTCCCAAGTTACCTCGGAATTAACATCACTGACACAGGGTTACCTTCGATCCCTGAGT	360
Db	313	AGTCCCAAGTTACCTCGGAATTAACATCACTGACACAGGGTTACCTTCGATCCCTGAGT	372
QY	361	TCTTAATCTTCTCTGAGATTTCCCTATGTATACACATCCGTGTTTTCCTTAAGTACTCA	420
Db	373	TCTTAATCTTCTCTGAGATTTCCCTATGTATACACATCCGTGTTTTCCTTAAGTACTCA	432
QY	421	GTTGAATGTTCTCTCTCTCCCTCCAGGGGTTTCCGTTTGCCCTCCCTCAAGGTTTTT	480
Db	433	GTTGAATGTTCTCTCTCTCCCTCCAGGGGTTTCCGTTTGCCCTCCCTCAAGGTTTTT	492
QY	481	TTGACAGCTGACACACCCGCTGCCACCTATTGACGTGAGACCTGCTGACGTGACCC	540
Db	493	TTGACAGCTGACACACCCGCTGCCACCTATTGACGTGAGACCTGCTGACGTGACCC	552
QY	541	TCTTAACAGCCACACTGTAGACAGCTGAGCTGCTGTCAGAGGGCCCTGTGTCAGCTGAGCC	600
Db	553	TCTTAACAGCCACACTGTAGACAGCTGAGCTGCTGTCAGAGGGCCCTGTGTCAGCTGAGCC	612

Dy 601 TGCTGCAGAGGGCACCCTGTTGAAGCTGAAGCCCTGTG 635  
Db 613 TGTGTCAGAGGGCACCTGTTGAGGCTTGAAGCCCTGTG 647

RESULT 12	
MSM074459	
ID	MSM074459
	standard; RNA; EST; 672 BP

DT	09-MAY-2003 (Rel. 75, Created)
DT	09-MAY-2003 (Rel. 75, Last updated, Version 1)

DE Homo sapiens mRNA; EST DKFZp6886A18242\_r1 (from clone DKFZp6886A18242)  
XX  
KM EST; expressed sequence tag.  
XX

OS Homo sapiens (human)  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Buteleostomi; Mammalia  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.

RP 1-67

RA Ottenwaelder B., Obermaler B., Deutschenbaur S., Mewes H.W., Weill B.  
RA Amlid C., Osanger A., Fodor G., Han M., Wilemann S.;

Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases

CC This is the 5' sequence of the clone insert  
CC from S. Wiemann, Molecular Genome Analysis, German Cancer  
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de,  
CC sequenced by Medigolomix (Martinriedel/Germany) within the CDNA  
CC sequencing consortium of the German Genome Project.

CC This clone (DKFZp686A18242) is available at the RZPD in Berlin.

FH	Key	Location/Qualifiers
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

FT	source	1. .672
11		

```

/mol_type="mRNA"
/mol_name="mRNA"

```

```
FT /clone="DKFZp6866A18242"
```

```
PT DH10B; sites SE11A + SE11B
PT /dev stace="adulr"
```

```

XX      /class_type="cDNA-collection"
XX

```

sequence 0/2 BF; 150 A; 164 C; 120 G; 201 I; 3

Query matrix	00.38 /	Score 01.2 /	DD 2 /	Zhang et al. 0 / 2
Best Local Similarity	98.3% /	Pred. No. 2,3e-89 /		

Matches 649; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

QY	1	AGCAGGCTAAAAAGCAGTAATCTTGGTTCCTCTCTAAAAGAAAGAAAATATAATT	60
Db	13	AGCAGGCTAAAAAGCAGTAATCTTGGTTCCTCTCTAAAAGAAAGAAAATATAATT	72
QY	61	TAAAAATACATTCGTAATTTCTTAAAAACAATAAATTAATGTTAAATTCATAGGCTC	120
Db	73	TAAAAATACATTCGTAATTTCTTAAAAACAATAAATTAATGTTAAATTCATAGGCTC	132
QY	121	AATCAAAATGAACCTTCCTCTTGGGCGCAGTATGTAATGCTTATGGACAGGAAGAG	180
Db	133	AATCAAAATGAACCTTCCTCTTGGGCGCAGTATGTAATGCTTATGGACAGGAAGAG	192



KW EST; expressed sequence tag.  
 XX Homo sapiens (human)  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homidae; Homo.  
 XX  
 RN [1]  
 RP 1-618  
 RA Behr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fodor G.,  
 RA Han M., Wiemann S.;  
 RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY  
 XX  
 CC This is the 5' sequence of the clone insert  
 CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 CC sequenced by Olgem (Hilden/Germany) within the CDNA sequencing  
 CC consortium of the German Genome Project.  
 CC No s1 sequence available.  
 CC This clone (DKFZp686A01243) is available at the RZPD in Berlin.  
 CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,  
 CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de  
 XX  
 FH Key Location/Qualifiers  
 FT source 1. .618  
 FT /db\_xref="taxon:9606"  
 FT /mol\_type="mRNA"  
 FT /organism="Homo sapiens"  
 FT /clone\_id="DKFZp686A01243"  
 FT /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
 FT DH10B; sites SfiIA + SfiIB"  
 FT /dev\_stage="adult"  
 FT /tissue\_type="CDNA-collection"  
 XX  
 SQ Sequence 618 BP; 143 A; 166 C; 118 G; 191 T; 0 other;  
 Query Match 59.9%; Score 603.4; DB 2; Length 618;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-87;  
 Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGACAGACTTAAAGGACATGATCTTCTGTTCTCTTAAAGAAAGAAATTAATTT 60  
 DB 14 AGACAGACTTAAAGGACATGATCTTCTGTTCTCTTAAAGAAAGAAATTAATTT 73  
 QY 61 TAAATATACATGCGTATTTCTTAAACATTAATTAATGCTTAATCTATAGGCTC 120  
 DB 74 TAAATATACATGCGTATTTCTTAAACATTAATTAATGCTTAATCTATAGGCTC 133  
 QY 121 AATCAAAATGAACTTCTCTTGGGCTGATGATGTTGCTTTGCAAGAAAG 180  
 DB 134 AATCAAAATGAACTTCTCTTGGGCTGATGATGTTGCTTTGCAAGAAAG 193  
 QY 181 AGCGTTCCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 DB 194 AGCGTTCCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 253  
 QY 241 GAATATTCCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 254 GAATATTCCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 313  
 QY 301 AGTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 DB 314 AGTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373  
 QY 361 TCTTAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 DB 374 TCTTAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433  
 QY 421 GTTGAATGTTCT 480  
 DB 434 GTTGAATGTTCT 493

QY 481 TTCAGAGCTGACGACCCGCTGCCCCACCTATTGACGCTGACCTGACCTGACCC 540  
 DB 494 TTCAGAGCTGACGACCCGCTGCCCCACCTATTGACGCTGACCTGACCTGACCC 553  
 QY 541 TCTTAAGCCACACCTGTATGACGCTGACGCTGACGCTGACGCTGACGCTGACCC 600  
 DB 554 TCTTAAGCCACACCTGTATGACGCTGACGCTGACGCTGACGCTGACGCTGACCC 613  
 QY 601 TGCTG 605  
 DB 614 TGCTG 618  
 RESULT 15  
 ID HSM075337 standard; RNA; EST; 610 BP.  
 XX HSM075337  
 AC BX485150;  
 XX BX485150.1  
 SV  
 XX 09-MAY-2003 (Rel. 75, Created)  
 DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)  
 XX  
 DE Homo sapiens mRNA; EST DKFZp686A12246\_r1 (from clone DKFZp686A12246)  
 XX  
 KW EST; expressed sequence tag.  
 XX  
 OS Homo sapiens (human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homidae; Homo.  
 XX  
 RN [1]  
 RP 1-610  
 RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A., Fodor G.,  
 RA Han M., Wiemann S.;  
 RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY  
 XX  
 CC This is the 5' sequence of the clone insert  
 CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 CC sequenced by BMFZ (Biomedical Research Center at the  
 CC Heinrich-Heine-University, Duesseldorf/Germany) within the CDNA  
 CC sequencing consortium of the German Genome Project.  
 CC No s1 sequence available.  
 CC This clone (DKFZp686A12246) is available at the RZPD in Berlin.  
 CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,  
 CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de  
 XX  
 FH Key Location/Qualifiers  
 FT source 1. .610  
 FT /db\_xref="taxon:9606"  
 FT /mol\_type="mRNA"  
 FT /organism="Homo sapiens"  
 FT /clone\_id="DKFZp686A12246"  
 FT /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
 FT DH10B; sites SfiIA + SfiIB"  
 FT /dev\_stage="adult"  
 FT /tissue\_type="CDNA-collection"  
 XX  
 SQ Sequence 610 BP; 143 A; 164 C; 115 G; 188 T; 0 other;  
 Query Match 58.1%; Score 585.4; DB 2; Length 610;  
 Best Local Similarity 99.8%; Pred. No. 1e-84;  
 Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 11 AAAAGCAGTATGATCTTCTCTCTTAAAGAAAGAAATTAATTAATTAATTAATTA 70  
 DB 24 AAAAGCAGTATGATCTTCTCTCTTAAAGAAAGAAATTAATTAATTAATTAATTA 83



```
QY 71 TTGGTATTTCTAAACATTAATTTATAGTGAATATTCATAGGTCATCAAAATG 130
Db 84 TTGGTATTTCTAAACATTAATTTATAGTGAATATTCATAGGTCATCAAAATG 143
QY 131 AAGCTTCTCTTGGGCGCTGCAATGATGTGTGCTTTGCAAGGAGAGACGGTTCCC 190
Db 144 AAGCTTCTCTTGGGCGCTGCAATGATGTGTGCTTTGCAAGGAGAGACGGTTCCC 203
QY 191 TTCAATTGGAGGATGACATATGATGTCACCCACTTCATCCATCTCTGAAATTTCC 250
Db 204 TTCATTGGAGATGACATATGATGTCACCCACTTCATCCATCTCTGAAATTTCC 263
QY 251 TATGGCATPACGGATTTACCACTCTCTTATATGCGCCAGATACAGTCCCACT 310
Db 264 TATGGCATPACGGATTTACCACTCTCTTATATGCGCCAGATACAGTCCCACT 323
QY 311 TACCTGGGAAATACTTACATGACACAGGGTTACCTGTATCCCTGATTCCTMACTTCT 370
Db 324 TACCTGGGAAATACTTACATGACACAGGGTTACCTGTATCCCTGATTCCTMACTTCT 383
QY 371 CCTGATTCCTTATGTCTATACATCCGTGTTTCCCTTAGCTACTCAGTTGAATGTT 430
Db 384 CCTGATTCCTTATGTCTATACATCCGTGTTTCCCTTAGCTACTCAGTTGAATGTT 443
QY 431 CCTCTCTCCCTCTAGGGGTTTCCCGTTGTCCCTCCTTGAAGGTTTTTTCAGACAGT 490
Db 444 CCTCTCTCCCTCTAGGGGTTTCCCGTTGTCCCTCCTTGAAGGTTTTTTCAGACAGT 503
QY 491 GCAGACCCGCTGCCCACTTATGACGTGAGCCTGCTGACGTGACACTTTCACAGCC 550
Db 504 GCAGACCCGCTGCCCACTTATGACGTGAGCCTGCTGACGTGACACTTTCACAGCC 563
QY 551 AACCTGTAGCAGTGAAGCTGCTGACAGGGGCCCTGTGACGCTGA 597
Db 564 AACCTGTAGCAGTGAAGCTGCTGACAGGGGCCCTGTGACGCTGA 610
```

Search completed: January 16, 2004, 09:22:47  
Job time : 2135 secs